Should Dying Patients Get Unproved Drugs? (*page 92*)

SCIENTIFIC AMERICAN

Saving Nature That WORKS FOR PEOPLE
page 50

October 2007 www.SciAm.com

SPECIAL REPORT

The Future of EXPLORING SPACE

- A Preview of the *Orion* Spaceship
- 5 Essential Things to Do Up There

Consciousness

Scientists Debate How Neurons Make Us Aware

Diamond Chips

Sparkling Processors Spin Quantum Logic

Diagnosis to Go

New Portable Microlabs

© 2007 SCIENTIFIC AMERICAN, INC.

FEATURES

SCIENTIFIC AMERICAN October 2007 ■ Volume 297 Number 4

SPECIAL REPORT

60 The Future of Space Exploration

By Steven Ashley and George Musser

The launch of the Soviet Sputnik satellite half a century ago inaugurated the Space Age. What comes next?

62 To the Moon and Beyond

By Charles Dingell, William A. Johns and Julie Kramer White

Humans are returning to the moon. This time the plan is to stay a while.

69 Five Essential Things to Do in Space

By George Musser

Planetary scientists have a quintet of goals for exploring the solar system.

ENVIRONMENT

50 Conservation for the People

By Peter Kareiva and Michelle Marvier

Pitting nature and biodiversity against people makes little sense. Many conservationists now argue that human health and well-being should be central to conservation efforts.

DEBATE

76 How Does Consciousness Happen?

By Christof Koch and Susan Greenfield

One of the greatest mysteries in science is how brain activity gives rise to subjective experience. Two leading neuroscientists compare their differing theories.

PHYSICS

84 The Diamond Age of Spintronics

By David D. Awschalom, Ryan Epstein and Ronald Hanson

Revolutionary electronic devices can harness the spins of electrons instead of their charge. Such devices might one day enable room-temperature quantum computers—made of diamond.

ON THE COVER

The sun peeks over the dark disk of Mars, as seen from the moon Phobos, in this illustration by artist and author Ron Miller.

© 2007 SCIENTIFIC AMERICAN, INC.

MORE FEATURES

MEDICINE

92 **Experimental Drugs on Trial**

by Beryl Lieff Benderly

A controversial lawsuit challenges the FDA's system of controlling access to experimental drugs and, some say, the scientific basis of drug approval.

INNOVATIONS

100 **Big Lab on a Tiny Chip**

By Charles Q. Choi

Squeezing a chemistry lab down to fingernail size could provide instant medical tests at home and on the battlefield.

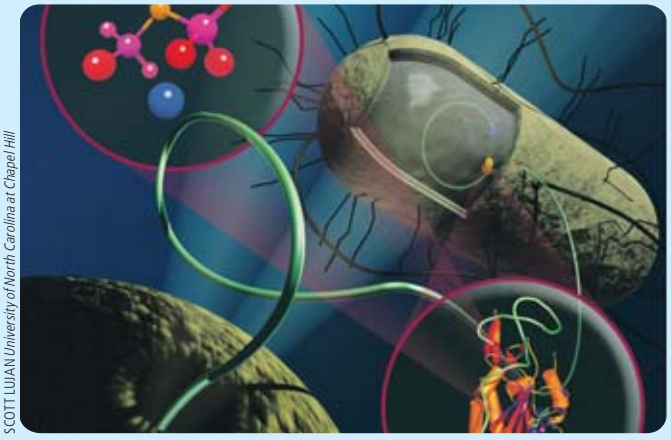
SCI AM

Yes or No?

Tell us whether patients should have the right to demand experimental medications. Go to **www.SciAm.com/ontheweb**

SCIAM.COM

GENES FOR ANTIBIOTIC RESISTANCE ▼ (*below*) can spread through a population of bacteria when individuals swap DNA. An already approved drug can stop this process and could help prolong the useful life of existing antibiotics.



SCOTT LUAN *University of North Carolina at Chapel Hill*

LIKE WHAT YOU SEE? Everything pictured here, plus Web-exclusive supplements to the articles in this issue, is available at **www.SciAm.com/ontheweb**

In Focus

Falcons, Ferrets and Forests Benefit from Preservation Efforts

Modern conservation projects deliver measurable effects, from rebounding ferret populations to less destructive logging in the Peruvian Amazon.

News

What Explains Toddlers' Linguistic Leap? Math

Simple math may explain why toddlers experience a sudden burst of words—and why some talk earlier.

Podcast

Reading Protects Brain from Some Toxic Exposure

Lead smelter workers who read more suffered less brain insult from lead exposure. Kevin Begos reports.

Blog

How Physics Embiggened

How a fake word from *The Simpsons* ended up in a perfectly cromulent string theory paper.

Strange but True

Infinity Comes in Different Sizes

If you were counting on infinity being absolute, your number's up.

WWW.RSPB-IMAGES.COM (*Eurasian bittern*); LURAY PARKER *Wyoming Game and Fish Department* (*ferret*); 20TH CENTURY FOX. ALL RIGHTS RESERVED/COURTESY EVERETT COLLECTION (*Homer Simpson*)

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111.

Copyright © 2007 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for public or private use, or by any information storage or retrieval system, without the prior written permission of the publisher.

Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

© 2007 SCIENTIFIC AMERICAN, INC.

DEPARTMENTS

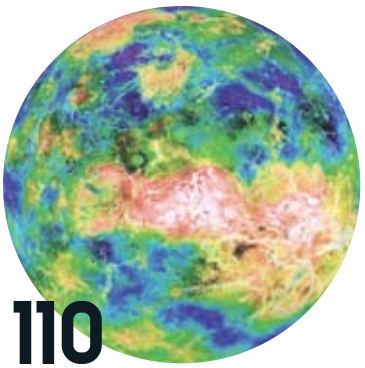
10 From the Editor

12 Letters

16 50, 100 and 150 Years Ago

18 Updates

38


110

News Scan

20

- Meeting more stringent mileage standards.
- An iron-seeded ocean to trap CO_2?
- Antiterrorism research benefits cancer patients.
- Time before the big bang.
- Zoom lens that doesn't need to zoom.
- Toxicity of fluorescent lightbulbs.
- Mixed messages in hormone therapy.
- Data Points: Dangerous space junk.

Opinion

40 **SciAm Perspectives**
Competition matters less than conquering space.

42 **Sustainable Developments**
By Jeffrey D. Sachs
Malaria, one of the world's worst killers, could be stopped soon.

44 **Skeptic**
By Michael Shermer
The most useful science blends data, theory and narrative.

48 **Anti Gravity**
By Steve Mirsky
Some strange science stories of recent vintage.

104

ESKO PETTAY

SCI AM

Darwin's Grandmothers

Read an expanded version of this month's Insights on evolution at **www.SciAm.com/ontheweb**

104 **Insights**
Deadbeat granddads, life-shortening sons and genetically bullying brothers—biologist Virpi Lummaa's studies reveal how evolutionary forces shape later generations.

108 **Working Knowledge**
Geothermal energy.

110 **Reviews**
Rapturous sociability.
Armageddon avoided.
The allure of Venus.

114 **Ask the Experts**
What is a "fictitious force"?
Why do apple slices turn brown?

116 **Fact or Fiction?**
Do helmets attract cars to cyclists?

10

114

116

© 2007 SCIENTIFIC AMERICAN, INC.

Choosing Targets

In space and on Earth, progress depends on making smart choices

She was a stray mongrel picked up off the streets of Moscow, and her handlers called her by several names before someone in the budding Soviet space program tagged her as Laika. On Halloween of 1957 she was bundled inside Sputnik 2. Three days later—just one month after Sputnik 1 was launched and started the space race—she became the first living creature in history to leave Earth. That distinction was tragically brief because Laika died only a few hours later, apparently from a combination of stress and overheating, but she lasted long enough to suggest that humans, too, might survive weightlessness and find a future in space. In our current celebrations of the Sputnik golden anniversaries, let us also salute the memory of a very good dog.

In the half a century since then, space programs have seen plenty of other epic triumphs and tragedies. For NASA, the past couple of decades have had an "it was the best of times, it was the worst of times" flavor, with fantastic high points such as the Cassini mission to Saturn and the treks of the Mars rovers but also nadirs such as the shocking losses of the *Challenger* and *Columbia* space shuttles. The manned space program in particular has often seemed listless, executing missions in support of an international space station whose raison d'être has grown shabbier every year.

That has begun to change. President George W. Bush set long-term goals in 2004 for returning to the moon and going to Mars; one may question the wisdom of those targets and their financial impact on science (*Scientific American* certainly has), but it is nonetheless good to see momentum carrying NASA toward something.

The Constellation program's transportation system, including the new *Orion* vehicle, is planned to make those coming forays in space possible. A team of NASA and Lockheed Martin engineers involved in developing those craft offers a peek at them in "To the Moon and Beyond," which makes up one half of this month's special report on the future of space exploration, beginning on page 62.

In the other half, "Five Essential Things to Do in Space," on page 69, staff editor George Musser weighs what NASA and its counterparts in other nations ought to be doing. Out of the limitless possibilities for study in our solar system, what priorities should space scientists set? Readers may agree or disagree with his choices (and we encourage you to send your preferred list to editors@SciAm.com), but consensus is not the point. Attention to setting the right goals is the best way to keep the space program healthy for another 50 years.

There is also no shortage of crucial choices to make here on Earth about how best to protect imperiled nature. Any desirable future for this planet must be conducive to humans along with the rest of the biosphere. But how to achieve that?

Some past *Scientific American* articles have championed focusing protection efforts on biodiversity "hot spots." In "Conservation for the People," on page 50, however, conservationists Peter Kareiva and Michelle Marvier argue that more good—for both nature and people—will come from saving ecosystems that render valuable services to human communities. Perhaps the only way to convince society to invest more in conservation is by showing how much it already has at stake.

JOHN RENNIE
editor in chief

CONCEPT IMAGE of a future crew launch vehicle

FLYNN LARSEN (*Rennie*); NASA/MSFC

Among Our Contributors

CHRISTOF KOCH
studies consciousness and is a professor of cognitive and behavioral biology at the California Institute of Technology. His article in this issue is his fourth for *Scientific American*.

SUSAN GREENFIELD
is a neuroscientist, professor of pharmacology at the University of Oxford, and director of the Royal Institution of Great Britain, in addition to being a member of the British Parliament's House of Lords.

PETER KAREIVA
became chief scientist for the Nature Conservancy after working in academia for 20 years and spending three working on salmon conservation for the National Oceanic and Atmospheric Administration.

MICHELLE MARVIER
directs the Environmental Studies Institute at Santa Clara University. Her research focuses on conserving biodiversity and assessing the ecological risks of genetically modified organisms.

© 2007 SCIENTIFIC AMERICAN, INC.

SCIENTIFIC AMERICAN®
Established 1845

EDITOR IN CHIEF: John Rennie
EXECUTIVE EDITOR: Mariette DiChristina
MANAGING EDITOR: Ricki L. Rusting
CHIEF NEWS EDITOR: Philip M. Yam
SENIOR WRITER: Gary Stix
SENIOR EDITOR: Michelle Press
EDITORS: Mark Alpert, Steven Ashley, Graham P. Collins, Mark Fischetti, Steve Mirsky, George Musser, Christine Soares
CONTRIBUTING EDITORS: W. Wayt Gibbs, Marguerite Holloway, Michael Shermer, Sarah Simpson

EDITORIAL DIRECTOR, ONLINE: Kate Wong
NEWS EDITOR, ONLINE: Lisa Stein
ASSOCIATE EDITORS, ONLINE: David Biello, Christopher Mims
NEWS REPORTERS, ONLINE: JR Minkel, Nikhil Swaminathan
COMMUNITY EDITOR, ONLINE: Christie Nicholson

ART DIRECTOR: Edward Bell
ASSOCIATE ART DIRECTOR: Mark Clemens
ASSISTANT ART DIRECTOR: Johnny Johnson
PHOTOGRAPHY EDITOR: Emily Harrison
PRODUCTION EDITOR: Richard Hunt

COPY DIRECTOR: Maria-Christina Keller
COPY CHIEF: Daniel C. Schlenoff
COPY AND RESEARCH: Michael Battaglia, John Matson, Aaron Shattuck, Rachel Dvoskin, Smitha Alampur, Michelle Wright

EDITORIAL ADMINISTRATOR: Jacob Lasky
SENIOR SECRETARY: Maya Harty

ASSOCIATE PUBLISHER, PRODUCTION: William Sherman
MANUFACTURING MANAGER: Janet Cermak
ADVERTISING PRODUCTION MANAGER: Carl Cherebin
PREPRESS AND QUALITY MANAGER: Silvia De Santis
PRODUCTION MANAGER: Christina Hippeli
CUSTOM PUBLISHING MANAGER: Madelyn Keyes-Milch

BOARD OF ADVISERS

RITA R. COLWELL
Distinguished Professor, University of Maryland College Park and Johns Hopkins University Bloomberg School of Public Health

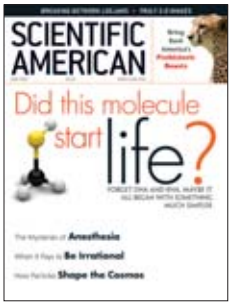
VINOD KHOSLA
Founder, Khosla Ventures

M. GRANGER MORGAN
Professor and Head of Engineering and Public Policy, Carnegie Mellon University

LISA RANDALL
Professor of Physics, Harvard University

GEORGE M. WHITESIDES
Professor of Chemistry, Harvard University

Rewilding ▪ Network Coding ▪ Traveler's Dilemma



JUNE 2007

"America's ecosystems have evolved since the Pleistocene; the consequences of introductions of exotic megafauna to the continent cannot be predicted."
—Dustin Rubenstein et al.

Rewilding Rebuttal

In "Restoring America's Big, Wild Animals," C. Josh Donlan reiterates a proposal to populate the American West with species (lions, cheetahs and elephants) he considers "proxies" for extinct megafauna present there in the Pleistocene. Donlan mentions some of our concerns, published in *Biological Conservation* in October 2006, but dismisses them unfairly. For example, we referenced *Jurassic Park* to emphasize his plan's sensationalism, not out of confusion about evolutionary timescales. There is a fundamental difference between "rewilding" and "Pleistocene rewilding." Rewilding involves reintroducing species extirpated within the past few hundred years to their native habitats; neither species nor habitats can have changed much in that time. The successful reintroduction of the Bolson tortoise to New Mexico is an example of this. Pleistocene rewilding would involve introducing exotic species Donlan hopes will fill the ecological roles of their Pleistocene ancestors.

What concern us are introductions of exotic megafauna where none have existed for thousands of generations. America's ecosystems have evolved since the Pleistocene; the consequences of such introductions cannot be predicted. We fear that threatened ecosystems will be harmed; new parasites and diseases will be introduced; the costs of fencing will be astronomical; and negative responses from local citizens could cause a conservation backlash. Pleistocene rewilding's proponents must marshal data to address these concerns. It will not be easy. The long life spans and vast home ranges of many large mammals may make conducting small-scale, controlled experiments impossible. Our views are not pessimistic but realistic. If Pleistocene rewilding is an "optimistic agenda," then its supporters must abandon sensationalism and begin outlining concrete plans for the future.

Dustin Rubenstein
University of California, Berkeley
Daniel Rubenstein
Princeton University
Paul Sherman
Cornell University
Tim Caro
University of California, Davis

Efficient Exchange?

"Breaking Network Logjams," by Michelle Effros, Ralf Koetter and Muriel Médard, presents a network-coding scenario in which a logjam potentially caused by two messages needing to travel one link between nodes E and F is avoided by sending evidence about those messages via the link and the messages via two unused links connecting to separate nodes.

Wouldn't it be more efficient to get rid of the unused links and add an additional link between E and F?

Barry Margolin
Arlington, Mass.

THE AUTHORS REPLY: *In transportation, it is not usually possible to remove an unused road to add lanes to one with heavy traffic. Similarly, rebuilding a communications network every time we encounter congestion is impractical. Especially so in wireless*

© 2007 SCIENTIFIC AMERICAN, INC.

SCIENTIFIC AMERICAN®
Established 1845

CHAIRMAN: Brian Napack
VICE PRESIDENT AND MANAGING DIRECTOR, INTERNATIONAL: Dean Sanderson
VICE PRESIDENT: Frances Newburg
CHAIRMAN EMERITUS: John J. Hanley

ASSOCIATE PUBLISHER, CIRCULATION: Simon Aronin
CIRCULATION DIRECTOR: Christian Dorbandt
RENEWALS MANAGER: Karen Singer
FULFILLMENT AND DISTRIBUTION MANAGER: Rosa Davis

VICE PRESIDENT AND PUBLISHER: Bruce Brandfon
SALES DEVELOPMENT MANAGER: David Tirpack
SALES REPRESENTATIVES: Jeffrey Crennan, Stephen Dudley, Stan Schmidt

ASSOCIATE PUBLISHER, STRATEGIC PLANNING: Laura Salant
PROMOTION MANAGER: Diane Schube
RESEARCH MANAGER: Aida Dadurian
PROMOTION DESIGN MANAGER: Nancy Mongelli

GENERAL MANAGER: Michael Florek
BUSINESS MANAGER: Marie Maher
MANAGER, ADVERTISING ACCOUNTING AND COORDINATION: Constance Holmes

DIRECTOR, SPECIAL PROJECTS: Barth David Schwartz

MANAGING DIRECTOR, ONLINE: Mina C. Lux
DIRECTOR, WEB TECHNOLOGIES, ONLINE: Vincent Ma
SALES REPRESENTATIVE, ONLINE: Gary Bronson

DIRECTOR, ANCILLARY PRODUCTS: Diane McGarvey
PERMISSIONS MANAGER: Linda Hertz

How to Contact Us

SUBSCRIPTIONS
For new subscriptions, renewals, gifts, payments, and changes of address: U.S. and Canada, 800-333-1199; outside North America, 515-248-7684 or www.SciAm.com

REPRINTS
To order reprints of articles: Reprint Department, Scientific American, 415 Madison Ave., New York, NY 10017-1111; 212-451-8877, fax: 212-355-0408; reprints@SciAm.com

PERMISSIONS
For permission to copy or reuse material: Permissions Department, Scientific American, 415 Madison Ave., New York, NY 10017-1111; www.SciAm.com/permissions or 212-451-8546 for procedures. Please allow three to six weeks for processing.

ADVERTISING
www.SciAm.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

networks, where user mobility leads to highly variable traffic patterns (an unused link may become a bottleneck later). For any network, network coding helps to maximize our rate of communication by making the best possible use of existing resources.

The Games We Assay

In "The Traveler's Dilemma," Kaushik Basu describes his game theory scenario in which two players each choose an integer from 2 to 100; the player choosing lower receives $2 more than that number and the other player $2 less. The game has a unique Nash equilibrium—an outcome from which neither can do better by deviating unilaterally—at which both choose 2. Yet in reality, higher choices often occur.

Such experiences do pose a serious challenge to the use of Nash equilibrium in game theory. But game theory is not synonymous with Nash equilibrium. There are now theorems giving conditions under which Nash equilibrium emerges. The assumption that players are rational is

DO YOU THINK I think you are rational?

far from sufficient to yield this outcome. There is no conflict between game theory and what we observe in Basu's game—only between Nash equilibrium and our observations. Basu asks for a "new kind of formal reasoning" that delivers more satisfactory analyses of many games. Over the past two decades, a new subfield of game theory—interactive epistemology—has tackled this topic. It is now possible to analyze mathematically what it means for players to be rational or irrational, to think that other players are rational or irrational, and so on. This approach is different from the classical Nash equilibrium analysis and often yields more intuitive answers.

Adam Brandenburger
New York University

Mortal Clay?

In "A Simpler Origin for Life," Robert Shapiro addresses life's origin but neglects a leading contender: the clay-based theory, which is even simpler than the metabolism-first theory he supports. In it, the first genome comes ready-made as clay crystals that replicate their patterns. The RNA-first theory he argues against is not the only possibility for a replicator-first origin.

Clay's affinity for organics entrains the use of simple organic molecules as aids in the competition for silica ions. These "naked genomes" evolve secondary, all-organic structures that take advantage of the geometry of some crystals. Genetic takeover complete, the new life-forms discard their clay structures for organic membranes.

A. K. Dewdney
University of Western Ontario

ERRATA "Breaking Network Logjams," by Michelle Effros, Ralf Koetter and Muriel Médard, lists Ning Cai, Shuo-Yen Robert Li and Raymond W. Yeung as affiliated with the University of Hong Kong. The Chinese University of Hong Kong is the correct affiliation.

"The (Other) Secret," by Michael Shermer [Skeptic], incorrectly refers to hydrogen's conversion to energy through nuclear fission, rather than nuclear fusion.

CLARIFICATION The June cover features methanethiol (CH_3SH), one of many molecules considered as possibly involved in a metabolism-first origin of life.

SCI AM

Read more on the Traveler's Dilemma and expanded versions of letters by Brandenburger and Rubenstein et al. at **www.SciAm.com/ontheweb**

Letters to the Editor

Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@SciAm.com

Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

MATT COLLINS

© 2007 SCIENTIFIC AMERICAN, INC.

"White Flight" ▪ Hypnotic Snakes ▪ Treasure Ship

Compiled by Daniel C. Schlenoff

OCTOBER 1957

METROPOLITAN SEGREGATION—"The white and non-white citizens of the U.S. are being sorted out in a new pattern of segregation. In each of the major urban centers the story is the same: the better-off white families are moving out of the central cities into the suburbs; the ranks of the poor who remain are being swelled by Negroes from the South. These population shifts bring with them profound economic consequences. Of first importance is a decline of parts of the central cities' business activity. In almost every city the big downtown department stores are losing trade to the suburban shopping centers. Retail sales in the central business district of Chicago fell 5 per cent between 1948 and 1954 while sales in the suburbs increased 53 per cent."

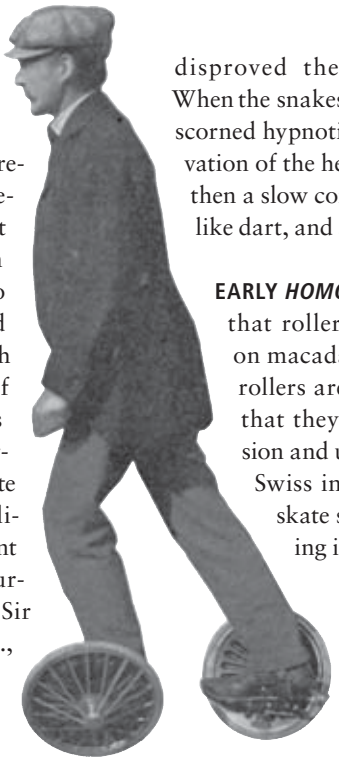
The entire article from 1957 is available at **www.SciAm.com/ontheweb**

PREMATURE—"Linear A has been cracked. Sixty years ago the British archaeologist Sir Arthur Evans unearthed some 2,000 baked clay tablets in the ruins of an ancient palace at Knossos on Crete. This strange ancient script found on them has long resisted decoding. Cyrus H. Gordon of Brandeis University, a specialist in Semitic languages and culture, has now deciphered Linear A and made the surprising discovery that it is the language of the ancient Babylonians, the so-called Akkadian language."

[EDITORS' NOTE: Gordon's work remains controversial; Linear A is considered to be as yet untranslated.]

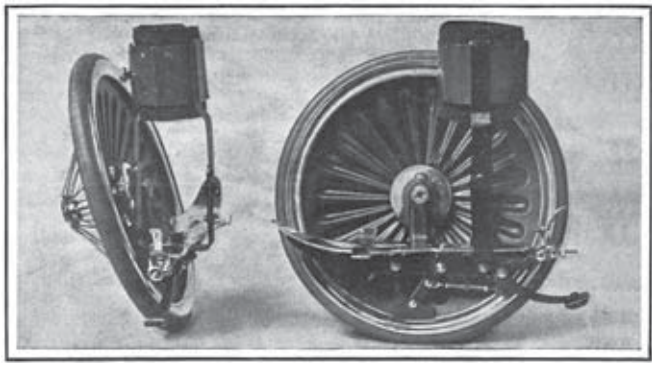
OCTOBER 1907

RAISIN FAN—"It is to be regretted that an economical and valuable article of food, in the shape of the dried currant, should be so much neglected. The dietetic value of the fruit is misunderstood and the prejudice against it entirely unjustified. Properly prepared, the currant might form an every-day item in the meals of the people, who seem to choose their food and arrange their dishes with an ignorance the extent of which is appalling. Let us make an interesting comparison, one that should be quite clear to the average intelligence: There is 54.87 per cent more total nutriment in currants than in lean beef.—Sir Francis Henry Laking, M.D., physician to the King of England [Edward VII]"



TRUSSST IN ME—"It is a popular belief that serpents have the power of capturing their prey by casting a mysterious spell over the victims. Even scientists have seriously considered this supposed mesmeric power over birds. Alfred Russel Wallace ascribed it to 'optic influences, akin to hypnotism.' In the rural districts, both of Europe and North America, bird-charming snakes are classed with such indisputable phenomena as fish-deluding anglers. However, my experiments with rattlesnakes last summer have disproved the snake-charm theory. When the snakes got ready to strike, they scorned hypnotic artifices. A gradual elevation of the head, a noiseless approach, then a slow contraction of coils, a snap-like dart, and a leisurely retreat."

EARLY *HOMO ROTALIS*—"The reason that roller skates cannot be used on macadam roads is because the rollers are of such small diameter that they drop into every depression and unevenness of the road. A Swiss inventor has designed the skate shown in the accompanying illustrations."



PERSONAL WHEELS for badly kept streets, 1907

OCTOBER 1857

GOLDEN SHIP—"It is stated by many of the *Central America*'s surviving passengers, that there were very few on the ship whose immediate wealth did not amount to hundreds, while numbers reckoned their gold by the thousands of dollars. The greater portion of the passengers were returning miners [from California]; some coming hither in hopes of a life of greater ease, and others to get their families and return to go to the land of gold. But as the storm continued to rage, less and less of gold was thought of, and when, on Saturday, it became evident that they were likely at any moment to be buried beneath the waves, the wealthy men divested themselves of their treasure belts and scattered the gold upon the cabin floors, telling those to take who dared to test its weight—as a few ounces might carry them to death."

[EDITORS' NOTE: Sunk off the coast of North Carolina, the wreck was located in 1987; three tons of gold have so far been recovered.] ■

SCIENTIFIC AMERICAN, VOL. XCVII, NO. 17, OCTOBER 26, 1907

© 2007 SCIENTIFIC AMERICAN, INC.

Helium Shortage ▪ Cytokine Storm ▪ Martian Dust Threat ▪ Novel HIV Drug

Edited by Philip Yam

Into Thin Air

Helium prices have doubled in the past five years. The high demand is not exactly coming from people with party balloons to fill. Rather helium cools the superconducting coils of magnetic resonance imaging (MRI) devices, and the sale of those machines has grown tremendously, driving the demand for helium up by 25 percent since 2003. In contrast, helium production has increased by only about half as much.

In 2006 the U.S. sold 23,000 metric tons of helium, which filled 71 percent of the world's helium needs (Algeria and Russia supplied most of the rest). At least one third of the U.S.'s contribution came from the federal helium reserve. Started in 1961, when helium was considered to be a crucial military and technological resource, the stockpile had grown by 1996 to 170,000 metric tons, stored mostly in porous rock beds in the Cliffside gas field near Amarillo, Tex. As part of an effort to privatize government programs, a 1996 act mandated the sale of all but 2,900 tons by 2015. As a result, according to a 2000 National Academy of Sciences report, the total U.S. helium resources will disappear by 2035—probably sooner, because of rising demand. "If within the next five years, new sources of helium are not brought to market, there will be a helium shortage" if demand continues to grow at current rates, says Joseph Peterson of the Bureau of Land Management, the agency that manages the reserve. Recycling of this rare and nonrenewable resource may need to improve greatly to prevent shortfalls. —*Sourish Basu*

Forget the Storm

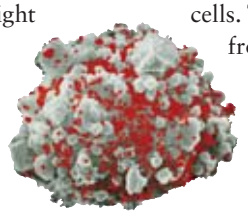
Researchers have hypothesized that taming the symptoms might boost survival odds during an influenza outbreak [see "Preparing for a Pandemic"; SciAm, November 2005]. Infections set off the release of cytokines, which are proteins that trigger inflammatory responses, including a rush of lymphocytes and the sacrifice of virus-compromised cells. The cytokine response to the avian flu virus, H5N1, is particularly vociferous, and some thought that this "cytokine storm" might be the main cause of death. Preventing it, unfortunately, does not appear to increase survival. Mice genetically engineered not to produce a key cytokine (TNF-alpha) fared no better than normal mice during infections. The same was true for ordinary mice given drugs to suppress cytokines.

The findings, in the July 24 *Proceedings of the National Academy of Sciences USA*, suggest that therapies should target the virus itself.

Choking on Mars

The Mars rovers Opportunity and Spirit seem to have weathered their biggest challenge yet—the weather. Throughout July especially, intense dust storms blocked sunlight from reaching their solar panels, causing severe power losses. At one point, Opportunity's energy production dipped to 128 watt-hours from a normal average of 700. Without enough juice, the onboard electronics could have frozen over, permanently disabling the craft. By mid-August the storms had abated enough to permit the machines to fully recharge their batteries. Considering the robotic explorers' remarkable toughness, they might be around to welcome the Phoenix lander. On August 4, NASA launched this craft, which will touch down on Mars on May 25, 2008. Its goal is to search for water on the planet's northern plains [see "The Red Planet's Watery Past"; SciAm, December 2006].

MODEL of Mars rover Opportunity.



HIV particles (*red*) bud from an infected white blood cell.

New HIV Fighter

In August, Pfizer got the go-ahead to sell Selzentry (maraviroc). The anti-HIV compound is the first of a class that works by blocking the CCR5 receptor, a primary portal through which the virus enters cells. The drug results from studies of people who have a resistance against HIV because they make mutated versions of CCR5 [see "In Search of AIDS-Resistance Genes"; SciAm, September 1997]. Perhaps up to 60 percent of HIV-positive people have the form of the AIDS virus that enters cells through CCR5.

CHARLES D. WINTERS *Photo Researchers, Inc. (helium tank)*; MARS EXPLORATION ROVER MISSION, CORNELL UNIVERSITY, JPL, NASA *(rover)*; TENGKU BAHAR *AFP/Getty Images (bird flu)*; NIBSC *Photo Researchers, Inc. (HIV particles)*

© 2007 SCIENTIFIC AMERICAN, INC.

AUTOMOTIVE TECH

Saving Gas and Lives

Can the U.S. improve fuel economy without sacrificing safety? BY MARK ALPERT

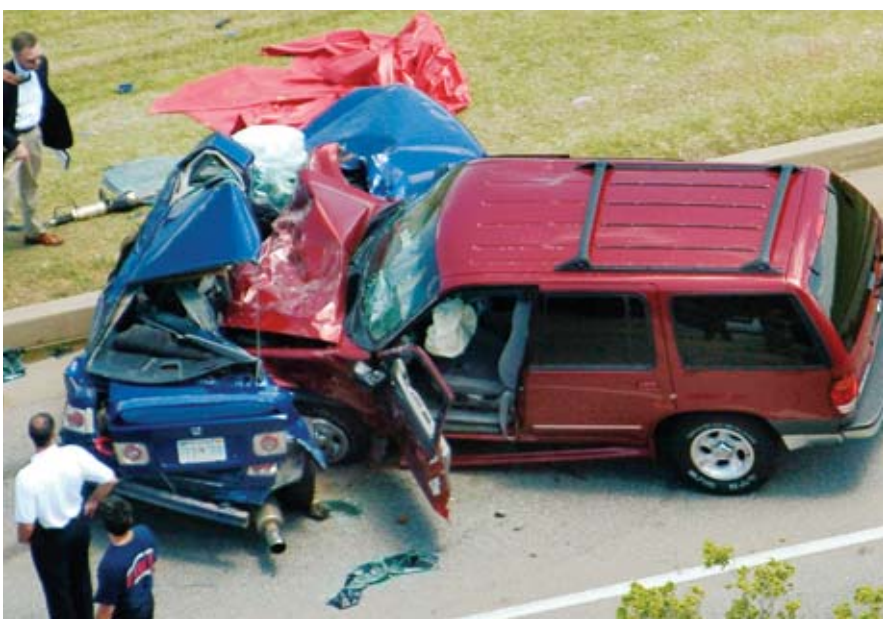
For years, the automobile industry has argued that congressional attempts to make cars and trucks more fuel-efficient would compromise passenger safety. The argument is based on the premise that the Corporate Average Fuel Economy (CAFE) standards imposed in 1975 resulted in a reduction of vehicle weights, which in turn caused about 2,000 traffic deaths a year that would not have occurred otherwise. But as Congress considers an energy bill that would further boost fuel economy—the combined average for cars and light trucks has been stalled at about 25 miles per gallon since the 1980s—transportation experts have disputed the contention that a lighter fleet would be less safe. What is more, new engine and transmission technologies could enable manufacturers to improve fuel efficiency without significantly cutting vehicle weights.

The current CAFE standards are 27.5 mpg for cars and 22.2 mpg for light trucks. In June the Senate passed a bill that would boost the fuel economy of new cars and trucks by about 40 percent. (In August the House of Representatives passed a bill that would leave CAFE standards unchanged; a House-Senate conference committee is expected to resolve the differences between the bills later this year.) Under the Senate proposal, the National Highway Traffic Safety Administration (NHTSA) would divide the fleet into classes based on size or weight and set fuel-economy standards for each class to achieve an overall average of 35 mpg by 2020. Tom Wenzel, a transportation scientist at Lawrence Berkeley National Laboratory, says the safety impact would depend on whether the regulations alleviate the present mismatch between cars and light trucks. According to Wenzel, the lower CAFE standard for trucks has fostered a proliferation of behemoth SUVs and pickups that cause thousands of deaths every year when they plow into cars.

A step in the right direction, Wenzel says, would be defining the vehicle classes by size rather than weight. Because the size of a vehicle's "crumple zone" can be crucially important for protecting passengers in a front-end crash, automakers should be discouraged from shrinking cars to enhance fuel economy. The best solution would be incorporating lighter, high-strength materials into auto frames and bodies, which would allow manufacturers to slash weight without trimming the vehicle's footprint.

Other experts note, however, that major weight reductions may not even be necessary. Says David Greene, a transportation researcher at Oak Ridge National Laboratory: "If manufacturers were to take the available technologies and apply them to fuel economy over the next 10 to 15 years, they could cost-effectively achieve a 40 to 50 percent improvement without making vehicles smaller." For example, some car engines already have variable valve lift and timing, which provides greater control over the flow of air into and out of the combustion chamber; until now automakers have employed this system primarily to boost horsepower, but it can also be used to increase fuel economy.

Another promising technology is a



STEVE GOOCH *The Oklahoman*/AP Photo

FATAL MISMATCH: The size discrepancy between SUVs and cars, fostered in part by differing fuel-economy standards, has increased the number of traffic deaths in the U.S. A careful adjustment of the standards could reduce the disparity in vehicle weights.

© 2007 SCIENTIFIC AMERICAN, INC.

camless engine, which opens and closes its valves by computer signals instead of connecting them mechanically, via cams and camshafts, to the crankshaft. Valeo, a French automotive company developing camless systems, estimates that the technology could raise fuel efficiency as much as 20 percent. Innovations in transmissions, aerodynamics, tires and power-hungry accessories such as air conditioners can also upgrade fuel economy.

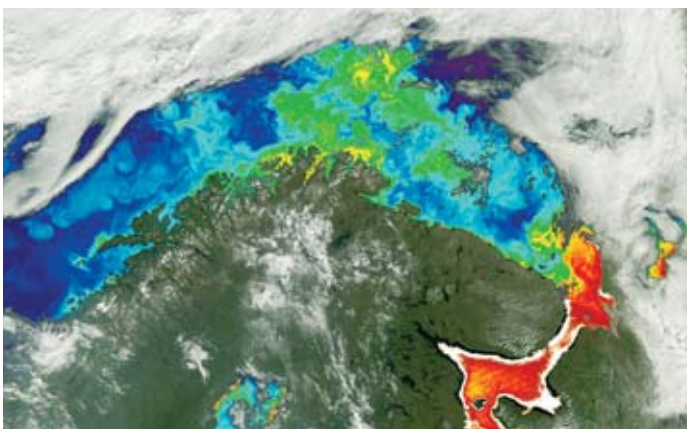
Greene predicts that the expense of meeting the new CAFE standards proposed in the Senate bill would probably range between $1,000 and $3,000 per vehicle, an up-front cost that could be recouped in fuel savings within five years if the price of gasoline remains above $3 a gallon. At the same time, the NHTSA could reduce the number of traffic fatalities by requiring the wider adoption of new safety features such as improved seat belts. "What we need are better safety regulations," Wenzel points out, "not inaction on CAFE."

CARBON SEQUESTRATION

Oceangoing Iron

A venture to profit from a CO_2-eating algae bloom riles scientists BY SOURISH BASU

Researchers have debated for a long time whether dumping iron into the ocean could ameliorate climate change. Iron encourages the bloom of tiny algae called phytoplankton, which take in carbon dioxide (CO_2) dissolved in the ocean for photosynthesis; that process in turn draws atmospheric CO_2 into the surface waters. Most scientists remain skeptical of whether iron fertilization will lead to greater carbon sequestration. But a company called Planktos, based in Foster City, Calif., has been forging ahead with such plans. Its latest target: 10,000 square kilometers of the equatorial Pacific, 600 kilometers west of the Galápagos—by far the most ambitious and controversial iron-seeding plan yet.



CARBON EATERS: Chlorophyll concentrations off the coast of Norway, depicted in false color from red (most dense) to purple (least dense), show a phytoplankton bloom in progress. The plankton on this page and the next are single-celled algae called diatoms.

Phytoplankton photosynthesize as much CO_2 as all the terrestrial plants combined. Although most of the fixed carbon returns to the ocean within a week when the phytoplankton die, perhaps up to a fifth of the biomass sinks to deeper waters, trapping carbon in the sea. For photosynthesis, phytoplankton need trace amounts of iron, and in the equatorial Pacific, the metal primarily comes from dust storms in the Gobi and Taklimakan deserts of Central Asia.

Citing a 2003 study by NASA and the National Oceanic and Atmospheric Administration (NOAA), Planktos claims that the amount of iron entering the equatorial Pacific has decreased by 15 percent since the early 1980s—presumably because of shifting wind patterns and fewer dust storms. As a result, the firm infers, phytoplankton populations there have decreased by 6 percent and CO_2 absorption by 3 percent. Through iron fertilization, Planktos wants to restore iron and phytoplankton to pre-1980 levels, which "will absorb 70 percent of the world's current CO_2 emission," affirms Planktos marketing chief William Coleman.

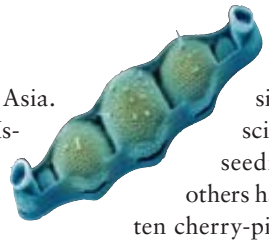
Dalhousie University oceanographer John Cullen, however, does not think that a low phytoplankton level necessarily indicates less carbon sequestration. A "successful" fertilization, he explains, will sink phytoplankton biomass—along with the nutrients it absorbs—to deeper waters, depleting the surface of chlorophyll (phytoplankton) and iron. Such declines might therefore be a signature of carbon being sequestered.

In any case, most investigators doubt that adding iron will shift the carbon load to the oceans. A dozen experiments have indeed shown that extra iron results in phytoplankton blooms, but whether carbon becomes trapped in the deep ocean over the long run remains unknown. In one such experiment, "less than 10 percent of the extra carbon that was fixed was actually making it to even 120 meters" of depth, says oceanographer Philip Boyd of the National Institute of Water and Atmospheric Research in New Zealand. And strong upwellings in the equatorial Pacific would probably release the trapped carbon back into the atmosphere in a few years anyway, says University of Southern California oceanographer Anthony Michaels.

Other data also counter Planktos's position that less

NORMAN KURING *MODIS/NASA (map)*; STEVE GSCHMEISSNER *Photo Researchers, Inc. (diatom)*

© 2007 SCIENTIFIC AMERICAN, INC.



iron has been blowing in from Asia. Air-sample analysis at Midway Island—the only station in the Pacific collecting in situ data on dust transport—"has not shown any declining trend [in Asian dust] from 1981 to 2001," says the University of Miami's Joseph Prospero, who is affiliated with the Midway observatory. Moreover, adds Texas A&M University atmospheric scientist Robert Duce, "there is no clear evidence of any shift in long-term wind pattern from the Asian highlands."

The low population of phytoplankton, if real [*see box at right*], may have nothing to do with the area's iron content. Natural decade-long climate cycles, such as El Niño and the Pacific Decadal Oscillation, could cause population variation, according to a 2006 study by marine algae specialist Michael Behrenfeld of Oregon State University and his colleagues.

"Any person who says that there's less iron in the oceans of the world than there was 100 years ago is speculating wildly, because we just don't have the data to make a call on that," in-

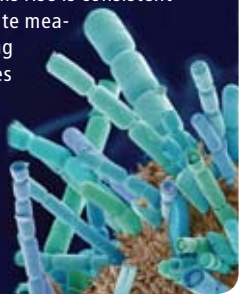
sists Boyd, who has been chief scientist on two of the 12 iron-seeding expeditions. Boyd and others have claimed that Planktos often cherry-picks data, and for a project that depends on scientific validation, the company has had trouble maintaining a panel of publicly identifiable scientists.

Even researchers who are sympathetic to iron fertilization, such as senior scientist Ken Johnson of the Monterey Bay Aquarium Research Institute, feel queasy about the Planktos project. "I wish there was a not-for-profit agency conducting this study," he says. Planktos has a financial incentive in that it plans to sell credits for carbon emissions based on the CO_2 the firm calculates will be sequestered by the phytoplankton blooms. Such credits will allow purchasers to emit CO_2 beyond agreed-on limits. On its Web site, Planktos even invites individuals to pay the company to "reduce" their carbon footprints.

Despite the criticisms—as well as protests from the U.S. government and environmental groups such as the global conservation organization WWF—Planktos remains undaunted. This past September it was to have poured about 50 tons of iron into the equatorial Pacific. In the absence of a global framework to verify the amount of carbon sequestered, Planktos may have seeded a lucrative, if scientifically tenuous, business.

Bloom or Bust?

In justifying its iron fertilization project in the equatorial Pacific Ocean, Planktos cites a 2003 NASA/NOAA report and infers a population drop in CO_2-consuming phytoplankton of 6 percent since the early 1980s. More recent studies, however, conclude just the opposite. David Antoine and his colleagues at the Oceanographic Laboratory of Villefranche in France reanalyzed the same data and found a 37 percent increase in chlorophyll levels (and hence phytoplankton). This rise is consistent with monthly on-site measurements spanning almost two decades at a station in the northern tropical Pacific, which show a 50 percent boost in phytoplankton population.



STEVE GSCHMEISSNER (*top and bottom*) AND SUSUMU NISHINAGA (*box*) *Photo Researchers, Inc.*

CANCER TREATMENT

Surviving Side Effects

Security fears spawn ways to treat radiotherapy's downside BY JOHN DUDLEY MILLER

After the September 11 attacks, Congress became worried that terrorists targeting the U.S. might explode a radiological weapon—most likely a "dirty" bomb, a kind of weapon that relies on a conventional explosive to spread radioactive materials packed around it. In 2004 Congress funded several research centers to create drugs to protect survivors and first responders from radiation injury. But the biggest beneficiary of this research might be a much different and far larger group of people: cancer patients.

Some 10.5 million Americans are living with cancer, according to the National Cancer Institute. These patients must conquer not one but two different diseases. "When we are talking about cancer survivors," explains Andrei Gudkov, senior vice president for basic research at the Roswell Park Cancer Institute in Buffalo, N.Y., "we mean survivors both from the disease itself and from the treatment of the disease." That is because the two common treatments—radiation therapy and chemotherapy—generally attack healthy tissue as well as tumors, causing long- and short-term complications. Radiotherapy sometimes even gives patients new tumors years later, Gudkov says. The complications also prevent many cancer patients from receiving doses large enough to treat the disease and survive. "So if we can reduce the chances of complications," says radiation oncologist Mitchell Anscher of Virginia Commonwealth University, "that's half the battle right there."

The 2004 legislation enabled the National Institute of Allergy and Infectious Diseases to create eight national Centers for Medical Countermeasures against Radiation in October 2005. The agency has spent $56 million in total so far, and it plans to commit another $82 million over the next three years.

© 2007 SCIENTIFIC AMERICAN, INC.

RADIATION THERAPY kills tumor cells but produces side effects that can limit its utility.

PAT SULLIVAN *AP Photo*

One strategy seeks to neutralize free radicals, which form because ionizing radiation knocks electrons off molecules, turning them into positively charged, highly toxic compounds. Researchers are working on drugs that mimic superoxide dismutase, a natural enzyme that transforms free radicals in the body back to harmless molecules before they can cause damage. Joel Greenberger of the University of Pittsburgh has shown in animals that manganese superoxide dismutase can protect the esophagus from damage during radiotherapy and up to 72 hours afterward. Phase I and II clinical trials testing the drug in patients with lung cancer are now under way. Anscher, Zeljko Vujaskovic of Duke University and others have shown that AEOL 10150, a small artificial molecule that mimics superoxide dismutase, protects against lung damage in rats.

Because so many more people now survive cancer over the long term than ever did before, researchers are also tackling one of the biggest problems that radiotherapy causes in these patients: fibrosis. Radiation oncologist Paul Okunieff of the University of Rochester calls this production of excess fiberlike connective tissues the most common side effect limiting cancer radiation doses to soft tissue. Fibrosis causes pain and swelling and restricts muscles' range of motion, lowering mobility and lung capacity. Anscher's group has focused on controlling TGF-beta, an immune system protein that causes cells to grow too much fibrotic tissue. Using antibodies and small molecules to attack TGF-beta before radiotherapy, his group has been able to reduce later fibrotic growth in mice and rats.

Existing cardiovascular drugs might also alleviate other long-term problems. Statins not only lower cholesterol but also protect blood vessels, which can be weakened by radiation. Researchers have begun human trials to see if lovastatin given before radiotherapy prevents rectal bleeding, which often occurs two to three years after prostate cancer treatment. ACE inhibitors, used to lower blood pressure, can reduce radiotherapy damage to kidneys, lungs and brains.

Some drugs could treat radiation effects from both therapy and bombs. For instance, Martin Hauer-Jensen of the University of Arkansas works with SOM230, a drug that mimics the hormone somatostatin. The drug inhibits the secretion of enzymes that destroy the inner lining of the intestines after radiotherapy. Because SOM230 can be given four hours after radiation, it should help in dirty bomb incidents. And the U.S. military seems especially excited about a drug from Cleveland Biolabs called Protectan CBLB502 to lessen radiation injuries in soldiers who have survived an atomic blast. According to Roswell Park's Gudkov, who is also Cleveland Biolabs's chief scientist, the drug targets a gene-activity regulator to boost production of superoxide dismutase, release immune cells that fight radiation damage, and limit apoptosis (cell suicide) in normal tissues—all properties useful in combating radiotherapy's side effects as well.

Amid post-9/11 concerns, many new ways to treat cancer radiotherapy's side effects will likely continue to emerge. As radiation oncologist John Moulder of the Medical College of Wisconsin puts it: "There's a lot more money now that it's called counterterrorism."

John Dudley Miller is based in Cleveland.

COSMOLOGY

New Beginnings

Ideas for a time before the big bang—which might be testable BY CHARLES Q. CHOI

The big bang is often thought of as the beginning of everything, including time, making any questions about what happened beforehand nonsensical. Now exotic theories that suggest the existence of an era *before* the big bang are growing in number. They indicate that imprints of this era might exist and that an upcoming generation of telescopes could detect them.

According to conventional big bang thinking, the universe emerged from a point of infinite energy and density, a singularity where the laws of physics break down. The universe then underwent "inflation," briefly expanding much faster than the speed of light. By smearing the cosmos out fairly evenly and smoothing out the early universe's curves, inflation solved a number of puzzles, including why spacetime is "flat," whereby light commonly travels in straight, not warped, lines. Ripples occurring during inflation could also explain the overall pattern, or structure, of galaxies seen now.

Observations of the cosmic microwave background radiation—the leftover heat from the big bang—have confirmed several broad predictions of the inflationary model. Still, inflation should have caused powerful gravitational waves that in turn should have distorted cosmic microwaves in detectable ways. The telescopes have not seen such distortions yet, ruling out

© 2007 SCIENTIFIC AMERICAN, INC.

several inflationary models. Moreover, critics say that the theories underlying inflation should mean that inflation is an eternal process; it should generate an infinite number of pockets of space with different properties, requiring more complex theories for why we live in a pocket that has the flatness and structure we see.

In the past 15 years, challenging theories arose that conjectured an era before the big bang, during which our universe contracted and then rebounded. Researchers say that the ekpyrotic scenario proposed in 2001 could successfully generate the current universe's structure, flatness and other features. (The name comes from the ancient Stoic notion of *ekpyrosis*, a fire in which the universe continuously gets reborn.) The cyclic model, derived from the ekpyrotic model in 2002, also accounts for the dark energy posited to be now causing universal expansion to accelerate [see "The Myth of the Beginning of Time," by Gabriele Veneziano; Scientific American, May 2004].

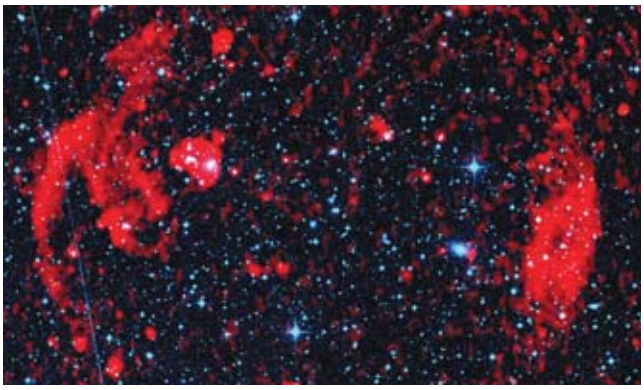
Still, these bouncing models did not convince many theorists. These scenarios posit that ripples before the big bang successfully passed the daunting barrier of a singularity to initiate structure in the current universe, an idea "most cosmologists are extremely skeptical of," admits Princeton University cosmologist Paul Steinhardt, who with University of Cambridge theoretical physicist Neil Turok helped to develop the ekpyrotic and cyclic models. In addition, the models were originally couched in terms of string theory, which many scientists disdain, because it calls for undetected extra dimensions of reality beyond those of space and time.

A flurry of new bouncing models has just burst out in the past few months. Strikingly, they come in a variety of different flavors, many of which avoid a singularity and all of which require no dimensions beyond those of space and time. "There's a lot of skepticism against

Inflation Deflation

Several new alternatives to cosmic inflation posit a cycle of birth and death for the universe. But not all alternatives demand such reincarnations. Robert Brandenberger of McGill University and his colleagues conjecture that the universe began with a hot, dense gas of strings, energy strands whose vibrations generate fundamental particles and forces. Thermal fluctuations of this gas then led to galaxy clusters and other cosmic structures. This model is "agnostic" as to whether anything existed before the big bang, Brandenberger says. It should, if correct, lead to gravitational signatures that future telescopes could detect.

© 2007 SCIENTIFIC AMERICAN, INC.



PRE–BIG BANG IDEAS must explain how large-scale structures such as clusters of galaxies (*blue*) formed.

JOYDEEP BAGCHI IUCAA, NRAO/AUI/NSF

bouncing, due perhaps to string theory," Steinhardt says. "These new results use more familiar physics and should convince most cosmologists—even those who don't want to consider extra dimensions—that there are real alternatives to inflation."

For instance, to prevent a singularity at the big bang, two models suggest that, essentially, a strong push kept the past universe from collapsing to a point. This force comes from a "ghost condensate," a fluid of exotic particles that can theoretically exert more pressure than even dark energy. These scenarios originated independently from theoretical physicist Burt Ovrut of the University of Pennsylvania and his colleagues and cosmologist Paolo Creminelli of the Abdus Salam International Center for Theoretical Physics in Italy, in partnership with Harvard University cosmologist Leonardo Senatore.

Another way to evade a singularity could be the intrinsic nature of spacetime. Relying on loop quantum gravity, an alternative to string theory, Pennsylvania State University theoretical physicist Martin Bojowald calculates that at extremely tiny scales, spacetime can become repulsive, preventing it from collapsing. A consequence of his scenario is what he calls "cosmic forgetfulness," in which the universe after the big bang forgets some of its past properties and acquires new ones independent of what it had before.

The new bouncing models should have resulted in post–big bang gravitational waves far weaker than inflation would generate, by 50 orders of magnitude. If more sensitive future telescopes, such as the Planck Surveyor, still fail to spot the distortions in the microwave background that inflation and its gravitational waves were supposed to have created, then such null results could support the idea of an era before the big bang. "At the moment I think it fair to say that inflation is more compelling," Creminelli says. "At the end, however, experimental data will decide between the alternatives."

Charles Q. Choi is a frequent contributor.

© 2007 SCIENTIFIC AMERICAN, INC.

Bright Horizons™
Holland America Line
TAMPA
GULF OF MEXICO
Key West
CARIBBEAN SEA
Costa Maya
Belize City
Santo Tomás de Castilla, Guatemala
Jan.27th–Feb.3rd, 2008
Western Caribbean
InSightCruises.com/SciAm
Cruise prices vary from $849 for an Inside Stateroom to $2,899 for a Full Suite, per person. (Cruise pricing is subject to change.) For those attending our Program, there is a $1,275 fee. Program subject to change. (Attendees may participate, at no charge, in the poker "lab" and seminars offered during the "Poker Passion" event: www.InSightCruises.com/Poker)
CO-PRODUCED BY:
SCIENTIFIC AMERICAN
InsightCruises
EDUCATION THAT TAKES YOU PLACES
CST# 2065380-40

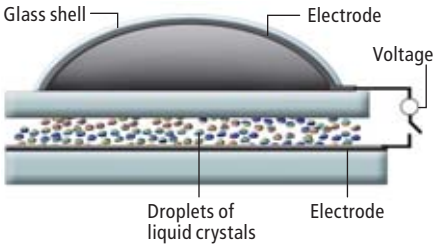
OPTICS

Liquid Zoom

Adaptive lenses change magnification without moving BY STEVEN ASHLEY

Camera bugs love zoom lenses, but they tend to be too bulky for cell phones and many miniature digital cameras. A research team at the University of Central Florida (U.C.F.) led by optics professors Shin-Tson Wu and Hongwen Ren has developed zoom lenses that can be dramatically smaller than conventional zooms. Whereas traditional zoom lenses move sets of lens elements mechanically to adjust focal length (and therefore magnification), the group's adaptive lenses alter focal length nearly instantaneously without changing the position of the lenses.

The U.C.F. team has two ways to make adaptive lenses. The first class, Wu explains, is based on the ability of a liquid-crystal layer to alter the degree to which it can refract (or bend) light when subjected to an electric field. In one version of this approach, the researchers adjust the layer's refractive index by varying the strength of the electric field in a concentric and graduated fashion. A change in the low voltage applied to a clear electrode modifies the focal length of the lens. Makers of cell phone lenses like to have a zoom magnification ratio of at least three and low voltages to save battery power, he says.



ADAPTIVE LENS uses a convex electrode that produces an electric field whose strength varies from center to edge. The light-refracting liquid crystals respond in kind, enabling changes in focal length (zooming).

The second approach "mimics the workings of the human eye," Wu states. It relies on a transparent fluid—water or oil—that is encapsulated between a see-through, flexible membrane and a flat glass substrate. When a tiny servomotor compresses a circular periphery seal that acts like an iris diaphragm, the shape of the membrane grows convex, modifying the fluidic lens's focal length.

The U.C.F. team has licensed five of its patented technologies to Holochip, a start-up company based in Albuquerque and San Francisco. It hopes to place the adaptive zooms, which can have apertures as small as a millimeter, into products sometime during the next few years.

ANN SANDERSON

ENVIRONMENT

Toxic Bulbs

Recycling rules vary for mercury-containing fluorescents BY DAVID APPELL

More consumers are placing compact fluorescent lightbulbs (CFLs) in their shopping baskets. Using 25 percent the energy of standard incandescents (and lasting 10 times longer), the swirly little tubes have become a symbol of green living and a means to fight climate change. Australia will require homeowners and businesses to replace all incandescents with CFLs by 2010, ultimately reducing greenhouse gas emissions by four million metric tons a year. At least four U.S.

© 2007 SCIENTIFIC AMERICAN, INC.

states and Congress are considering similar legislation.

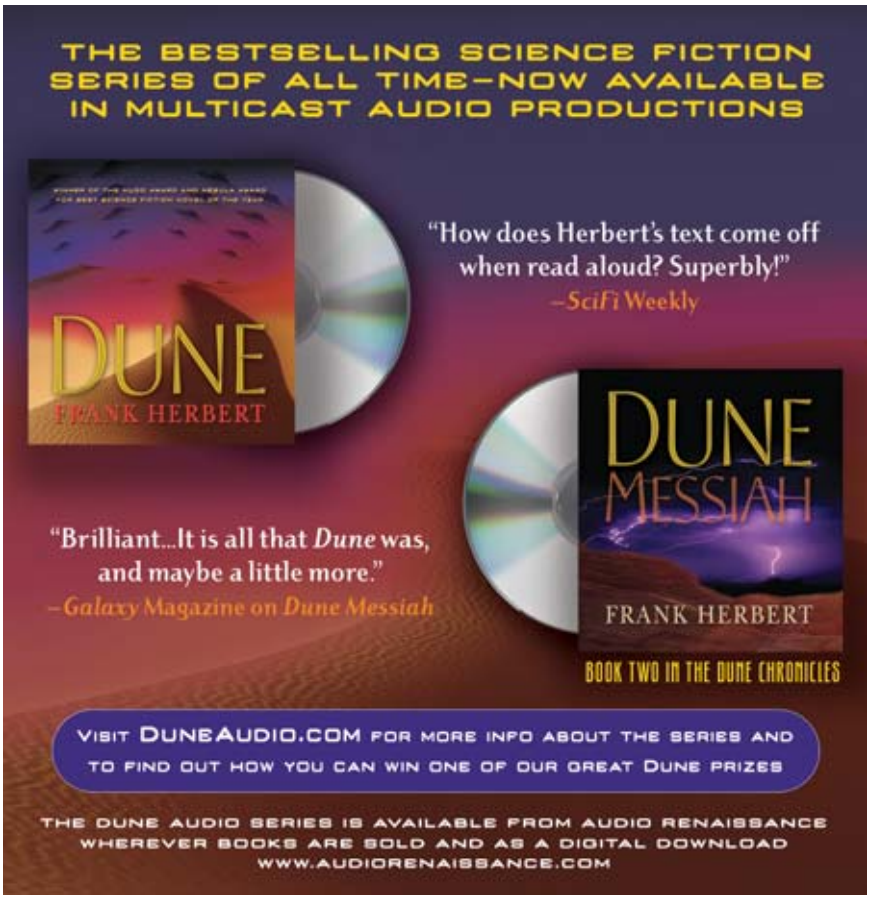
Yet CFLs have a downside: the bulbs contain mercury and cannot be tossed out with the ordinary trash. Roughly two billion will be sold in the U.S. this year (about 5 percent of the total lightbulb sales)—raising questions of how to handle 10 metric tons of mercury each year after the bulbs burn out.

Mercury is essential to the function of fluorescent bulbs. An electrostatic charge vaporizes the mercury and induces it to emit ultraviolet light, which makes the phosphorous coating inside the bulbs glow. A potent neurotoxin, mercury is especially dangerous for fetuses and children. In the U.S. today about one sixth of children are born having been exposed to mercury levels so high they are at risk for memory loss and learning disabilities, according to the Environmental Protection Agency.

Each CFL contains about five milli-

EDWARD KINSMAN *Photo Researchers, Inc.*

FIXTURE IN A GREEN HOME: Compact fluorescent bulbs use less energy than standard incandescent bulbs, but because they contain mercury, they pose an environmental risk of their own.



SCIENTIFIC AMERICAN

Subscriber alert!

Scientific American has been made aware that some subscribers have received notifications/subscription offers from companies such as United Publishers Network, Global Publication Services, Publishers Access Service, Lake Shore Publishers Service, Publishers Consolidated, Platinum Publishing Service and American Consumer Publishing Association. These are not authorized representatives/agents of Scientific American. Please forward any correspondence you may receive from these companies to:

Simon Aronin
Scientific American
415 Madison Ave.
New York, NY 10017

AKZO NOBEL

The Editors and Publishers of Scientific American wish to express our gratitude to Akzo Nobel for partnering with us to bring 750 new subscriptions to honors students in the following New York State High Schools and Competitions:

Ardsley High School
Brewster High School
Fox Lane High School, Bedford
Hastings-on-Hudson High School
Kingston High School
Mahopac High School
New Rochelle High School
Ossining High School
Rye High School
Saunders High School, Yonkers
Sleepy Hollow High School
The Westchester Science & Engineering Fair
Woodlands High School, Hartsdale

We wish you all the best of luck in this new school year!

SCIENTIFIC AMERICAN Education Outreach

© 2007 SCIENTIFIC AMERICAN, INC.

grams of mercury, about equal to the amount of ink on the tip of a ballpoint pen. Of course, mercury in a CFL does not pose the same kind of risk as, for instance, mercury in fish (the U.S. Food and Drug Administration sets a limit of 0.17 milligram in a six-ounce can of tuna). But it can leach out of landfills into water supplies or become airborne if incinerated.

Despite years of effort, manufacturers have failed to find a replacement for mercury, although they have succeeded in reducing the amount of mercury per bulb. "Manufacturers have grown more obvious in their attention to the mercury content of lamps," says Bill Stanwood of the Product Stewardship Institute, a Boston-based nonprofit that seeks to reduce the health and environmental impact from consumer goods. Whereas industrial users are familiar with the need to recycle linear tubes, residential consumers have yet to catch on—the recycling rate for fluorescent bulbs in the U.S. is about 24 percent, according to the Association of Lighting and Mercury Recyclers. "About one third of the country lives where you can't throw this stuff into the garbage legally," says the association's Paul Abernathy.

Currently CFL recycling options vary across the country. Sylvania offers a mail-back recycling kit that costs about $1 a bulb. Wal-Mart, which last year announced a goal of selling 100 million CFLs annually, now has kiosks for spent CFLs, but only in its California stores. The U.S. Postal Service is considering a recovery program with recycling containers at their stations.

At least one state is showing that CFL recycling is economically feasible. Vermont has one of the highest levels of CFL sales per household, and in 1998 it was the second state (after Minnesota) to pass a law requiring the recycling of CFLs. In August 2005 True Value hardware stores in Vermont began taking back customer's spent bulbs and shipping them back to warehouses on merchandise-delivering trucks. This "reverse distribution" process costs about 35 cents per bulb, says Karen Knaebel, Mercury Education and Reduction Coordinator for Vermont's Department of Environmental Conservation. (A state survey found that two thirds of Vermonters would pay 50 cents to recycle a bulb.)

With this strategy, Vermont has recycled 13,000 linear feet of traditional fluorescent lighting and 4,000 CFLs in almost two years. The recycling rate is currently increasing at 17 percent a year.

Stanwood's organization hopes to translate that kind of success to the rest of the country. His group is working on a plan to facilitate a national dialogue among interested parties to establish more standardized procedures for CFL recycling. But until they are set, consumers who want to recycle will have to fend for themselves once the lights go out.

David Appell lives in Portland, Ore.

And If You Break a CFL ...

Although compact fluorescent lightbulbs contain mercury, breaking one in your home will not require a costly visit by a hazmat team. Open windows to dissipate mercury vapor. Then, while wearing gloves, use sticky tape to pick up the small pieces and powdery residue from the bulb's interior. Place the tape and large pieces of the bulb in a plastic bag. After vacuuming the area, place the vacuum bag inside doubly sealed plastic bags before discarding.

Check the Environmental Protection Agency's Energy Star Web site, www.energystar.gov/index.cfm?c=cfls.pr_cfls, for more information.

To find a CFL recycling program in your area, go to www.lamprecycle.org

HEALTH

Easing Hormone Anxiety

For women just past menopause, hormone pills seem safe BY TABITHA M. POWLEDGE

Five years ago panicked postmenopausal women threw away their hormone pills after the federal government revealed that the drugs raised the risk of breast cancer and coronary disease. In 2006 only six million U.S. women got hormone prescriptions, a dramatic drop from 16 million in 2001. It seemed like the end for hormone therapy, especially for the previously fashionable notion that hormones protected older women from cardiovascular disease and other ills caused by aging.

But papers published this year, written by researchers who delivered the original warning, show that taking hormones soon after menopause—within about 10 years—is safe for most women. The data even suggest that hormone therapy for less than 10 years may improve some women's health more than doing nothing.

The uproar over hormone therapy began in July 2002, when researchers associated with the massive Women's Health Initiative (WHI) reported on one section that was following nearly 17,000 participants. Women taking a specific combination of estrogens and synthetic progesterone, they said, were experiencing small increases in breast cancer. WHI investigators also found bump-ups in stroke, pulmonary embolism and coronary heart disease. The National Institutes of Health,

© 2007 SCIENTIFIC AMERICAN, INC.

BY PATCH OR PILL, women can relieve symptoms of menopause with hormone therapy. Many fled from it in 2002 because of a report of increased breast cancer and coronary risk. A reanalysis suggests that the risk depends on age and the time since menopause.

which runs the WHI, sent warning letters to study subjects and abruptly shut down the trial. (Epidemiology studies reported this year have observed a drop in breast cancer rates since the trial ended, although some evidence suggests that the slide actually began before the termination.)

In 2004, however, the outlook for hormones began to improve. The WHI published new clinical trial data on hysterectomized women who took only estrogen. (Because estrogen can promote uterine cancer, women who have uteruses often also take some form of progesterone, which protects against the disease.) The estrogen trial also stopped early, this time because of slight stroke increases. But startlingly, these women suffered fewer heart attacks, and even less breast cancer, than women taking a placebo. Among 3,310 women ages 50 to 59 divided roughly equally between estrogen and placebo groups, 16 estrogen users developed coronary heart disease, compared with 29 in the placebo group; 35 placebo users developed breast cancer, compared with just 25 in the estrogen group.

That age strongly affects the outcomes received confirmation earlier this year, when WHI researchers reanalyzed their original data and combined them with the estrogen numbers. Besides generally verifying the decreased risk in heart disease, they also show that the response to hormone therapy depends on a woman's age and years since menopause. Even the inevitable rise in stroke and blood clots was less in younger women (those younger than 70). Women in their 70s, however, suffered similar heart disease rates whether they took estrogen or a placebo.

In looking back at the fallout from the 2002 report, Jacques Rossouw, who heads the WHI trials, acknowledges that women younger than 60 who would have been appropriate candidates for hormone therapy—those suffering severe hot flashes, for instance—fled from hormones along with older women. "My surmise is that women just got scared of hormone therapy across the board, irrespective of what they were using it for," he says. "With hindsight you could say, well, maybe we should have emphasized reasonable use even more." The study's abrupt termination may have also stoked fears. "Maybe we didn't need to do it that way," remarks WHI investigator Marcia Stefanick of Stanford University. "It wasn't an emergency—it wasn't like people were, you know, under serious threat of the adverse outcomes."

Wulf Utian, executive director of the North American Menopause Society, criticizes the WHI researchers because, as he puts it, "they've always taken the glass-half-empty read on their data." Utian also says that, especially between 2002 and 2004, declarations made in press releases or interviews were much more negative than the conclusions in published scientific papers. The data seemed particularly alarming because the risks often appeared in relative terms, such as a 29 percent increase for heart attacks between hormone users and nonusers, rather than in terms of total risk, in which the overall heart attack risk rose by just 0.07 percent, from 33 to 40 per 10,000 women per year.

JoAnn Manson, a WHI investigator at Harvard University, says that in 2002, analyses on years since menopause had not yet been done, and it took the 2004 estrogen-only results to make clear that younger women were at much less risk. Still, she notes, strong hints of an age effect in earlier clinical and animal studies existed: "Taking all of the previous research into account, there may have been a reason to look very closely at differences by age and differences by time since menopause." Had that been part of the earliest reports, Manson remarks, it might have helped put the results into perspective for younger women.

Overall, WHI researchers say they are pleased to have derailed the trend of prescribing hormones for women long past menopause. Still, Stefanick says, "I wish we had figured out a way to change prescribing practices but have fewer people distressed about it."

Tabitha M. Powledge is based near Washington, D.C.

IAN WALDIE *Getty Images (top);* JULES SELMES AND DEBI TRELOAR *Dorling Kindersley (bottom)*

© 2007 SCIENTIFIC AMERICAN, INC.

Data Points

Hit or Miss

On July 23, astronauts onboard the International Space Station (ISS) heaved out a 630-kilogram, refrigerator-size tank of ammonia and a 96-kilogram obsolete camera mounting. NASA frowns on littering in space but made an exception in this case because the pieces were too large to bring back on a shuttle mission (they will fall out of orbit in about a year and burn up). More dangerous are objects smaller than 10 centimeters wide, which are not tracked by ground stations. A one-millimeter-wide meteoroid can puncture an astronaut's suit. We asked Bill Cooke of the NASA Marshall Space Flight Center in Huntsville, Ala., and Mark Matney and Eric Christiansen of the NASA Johnson Space Center in Houston—all specialists in meteoroid and orbital debris impacts—to provide details about the possibility of collisions. ***—Sourish Basu***

Impact speed in kilometers per second of:

Meteoroids: 12 to 72

Debris: 5 to 15

Bullet from a handgun: 0.34

Size that is of concern:
> 1 millimeter

Main altitude of circling debris in kilometers: 700 to 1,000

Altitude of ISS and space shuttle in kilometers: 300 to 400

Percent chance of impact with:

Space debris: 0.12 to 0.2

Micrometeoroid: 0.08 to 0.13

Chance of losing a shuttle from an impact: 1 in 500 to 1 in 300

Metric tons per day of meteoroids entering Earth: 16.5

BIOLOGY

When Clean Living Isn't Longer Living

Even harmless bacteria suck up the energy of their hosts and hasten their deaths—or so the conventional thinking went. New findings show that flies scrubbed clean of bacteria do not outlive their infested brethren. In both flies and humans, because the number of bacteria living both inside and on the body increases with age, researchers expected infestations to prove harmful by depleting their hosts' resources. University of Southern California scientists and their colleagues compared normal fruit flies with ones born from eggs washed in antibiotics and raised in bacteria-free environments. The flies even ate disinfected food. Surprisingly, normal and superclean fruit flies had the same life span of roughly 65 days. Although these experiments cannot be replicated in higher organisms, which need bacteria for proper digestion and other functions, the investigators say their results do help narrow down which factors help to limit animal longevity. The findings appear in the August 8 *Cell Metabolism*. *—Charles Q. Choi*

PHYSICS

Paddle-Free Swimming

For microscopic swimmers, paddling is hard because at such tiny lengths, water seems as thick as honey. To get around this viscosity, scientists from the University of Sheffield in England and their colleagues have developed plastic swimmers without moving parts. They coated 1.6-micron-wide balls on one side with a thin layer of platinum and dropped them in a solution of hydrogen peroxide and water. The metal catalyzes the breakdown of hydrogen peroxide into oxygen and water, propelling the balls at speeds up to five microns per second, half as fast as similarly sized bacteria move. After a few seconds, the balls start to jostle with surrounding atoms and molecules and then meander, but the designers suggest that magnetic fields could guide the swimmers. They could also adapt them to work in fluids such as blood, to help deliver drugs within the body. Wade into the findings in the July 27 *Physical Review Letters*. *—Charles Q. Choi*

BEHAVIOR

Hot Tails from the Squirrel

Squirrels can heat their bushy tails to warn off infrared-sensitive snakes hungry for their pups. Infrared video showed that the tails of adult California ground squirrels, which are fearsome defenders able to resist snake venom, warmed by several degrees when threatened by rattlesnakes, which detect infrared via so-called pit organs in their noses. But no heating occurred while the rodents confronted gopher snakes, which lack such heat seekers. A stuffed squirrel also elicited defensive reactions from rattlesnakes if the mock critter's tail was artificially heated and wagged. The results, published online August 17 by the *Proceedings of the National Academy of Sciences USA*, mark the first discovery of one animal communicating with another by infrared. Still, the defense is not perfectly effective: earlier studies have found that a rattlesnake's diet is 70 percent squirrel pup. *—JR Minkel*

GEORGE McCARTHY *Nature Picture Library (squirrel)*; ILLUSTRATION BY MATT COLLINS

© 2007 SCIENTIFIC AMERICAN, INC.

In Brief

"VIRGIN BIRTH" STEM CELLS

The stem cells claimed to be extracted from the first cloned human embryo by discredited scientist Woo Suk Hwang actually owe their existence to parthenogenesis, in which egg cells give rise to embryos without being fertilized by sperm. A series of genetic markers sprinkled throughout the cells' chromosomes shows the same pattern found in parthenogenetic mice as opposed to cloned mice. The result, published online August 2 by *Cell Stem Cell,* suggests that Hwang and his group were the first to achieve human parthenogenesis but failed to recognize it. (International Stem Cell in Oceanside, Calif., announced successful human parthenogenesis this past July.) Parthenogenesis may offer a way of creating cells that are genetically matched to a woman for disease treatment. —*JR Minkel*

SUSPENSIONS FOR SAFER ROADS

Nearly 40 percent of all fatal traffic accidents in the U.S. involve drinking, resulting in 17,000 annual deaths. A surefire way to prevent some of these accidents: take away on the spot the license of anyone who fails a breath test. Investigators reporting in the August *Alcoholism: Clinical and Experimental Research* describe how they analyzed 26 years of federal crash data from 46 states with driver's license suspension laws and confirmed that immediate suspension reduced the number of such fatalities by 5 percent. Legal scholars note, however, that suspensions without due process may violate civil rights. —*Coco Ballantyne*

HOW TO ACT LIKE A MALE

Female lab mice tend to be docile, passive creatures. But by either genetically shutting down or surgically removing their ability to smell pheromones, scientists transformed them into aggressive, pelvic-thrusting, vocalizing lotharios—without any significant rise in male hormones. The key to the behavioral change is a cluster of receptors in their nose called the vomeronasal organ, which connects to the brain and registers the gender of other mice, triggering the appropriate response. The study, published online August 5 by *Nature,* suggests that gender-specific behaviors may have less to do with hormones and more to do with how neural circuitry gets triggered. —*David Biello*

PHYSICAL ANTHROPOLOGY

Cohabitating Hominids

The human ancestral family tree just got a lot bushier: a newly discovered fossil supports the notion that *Homo habilis* may not have evolved into the larger *Homo erectus*. Researchers have found an *H. habilis* jawbone near Kenya's Lake Turkana that dates to 1.44 million years ago, when *H. erectus* also roamed the area. If the two early hominids coexisted, then they probably evolved separately from a common ancestor. The oldest *habilis* and *erectus* fossils found in East Africa date to 1.9 million years ago, indicating that they cohabited in the region for half a million years and that the species' shared ancestor would have had to have lived two million to three million years ago. But whether it was much different from both or resembled the older *habilis* is unknown: the only remains from this time are some skull fragments and a few teeth. Dig up the study in the August 9 *Nature*. —*JR Minkel*

FAMILY CONNECTIONS: *Homo habilis* may not have evolved into *H. erectus.* The *H. habilis* skull shown dates to 1.8 million years ago.

LINGUISTICS

Bursting with Words

Between the ages of one and two, toddlers seem to go from babbling hesitantly to confidently chatting up a storm. But how? Researchers have proposed that toddlers acquire a special function that enables this verbal explosion, such as syntactic bootstrapping (using a new word's context to determine its meaning). Bob McMurray of the University of Iowa devised a mathematical model to study the problem and found that tots inevitably experience a rapid increase in word learning after a set period. Initially little tykes can repeat only the most often heard words, but four months later they might command several hundred words that they hear only occasionally. This learning curve takes the shape of a simple, normal statistical distribution—a bell curve—and any child mastering language will travel along it. The curve may also explain why some children learn language faster than others: they might have been exposed to more words or have better short-term memory. —*David Biello*

PASCAL GOETGHELUCK *Photo Researchers, Inc. (top);* MANGO PRODUCTIONS/CORBIS *(bottom)*

Read More ...

News Scan stories with this icon have extended coverage on **www.SciAm.com/ontheweb**

© 2007 SCIENTIFIC AMERICAN, INC.

SciAm Perspectives

Racing past the Moon

Today competition matters less than conquering space

BY THE EDITORS

Fifty years ago the starting pistol fired with the launch of Sputnik, and the space race between the Soviet Union and the U.S. was on. The contest was less like a race, however, than it was an extraplanetary game of Risk, with each side seeking to attain dominance over some portion of Earth's neighborhood. Notwithstanding the high-minded rhetoric about humanity expanding out among the stars, the leadership of the two countries had more immediate military motivations to develop better technology for hurling ICBMs and spying from orbit—and denying the other side an equivalent advantage.

Then, after the Apollo program succeeded in putting people on the moon, the political leadership failed to embrace inspiring new goals and the international competition collapsed from a lack of momentum. NASA not only withdrew from its greatest achievements but scuttled the spacecraft that enabled them. Unbelievable as it seems to those of us who watched Neil Armstrong in the Sea of Tranquility, no one under the age of 35 in 2007 can remember humans standing on another world.

For many space science enthusiasts, that situation is not even remotely troubling. The past decade and a half have been a golden era of discovery thanks to the Hubble, COBE and other space-based telescopes, as wells as robotic probes to Mars and the outer planets. Unmanned ventures have stumbled—witness the loss of the Mars Observer and Mars Polar Lander missions—but never to a degree rivaling the tragedies that darkened the space shuttle program or the fiascoes that have dogged the International Space Station. And the scientific return on investment has been incalculably greater. But it must be acknowledged that even the inspiring Mars rovers cannot quicken the pulse as Apollo did.

Recently space programs have suffered serious blows. As this issue goes to press, NASA is dealing with allegations that some astronauts were drunk at launch. Three workers for the California-based spacecraft maker Scaled Composites were killed in an explosion. Meanwhile RKK Energia, a principal manufacturer of Russian spacecraft, is flirting with bankruptcy. Terrible as these setbacks are, however, they do not eliminate the fundamental appeal of exploration.

Today's space race is far more complicated than the earlier one. Now the U.S. vies not only with Russia but also with the European Space Agency, China and other rising powers. Private enterprises such as Virgin Galactic and SpaceX are staking their own claims on chunks of the emerging space industry. All these space racers have set themselves somewhat different goals, from the wildly ambitious (voyaging to Mars or establishing a permanent base on the moon) to the straightforwardly commercial (cornering the satellite launch market or operating the first transatmospheric tour bus).

Space is big; there is room for all. Diverse spacefarers can complement and cooperate more successfully than they can stymie one another. Companies can and should take on ambitious ventures beyond Earth, but there are also worthwhile undertakings that only governments have the wherewithal to do in the near term, and they will lay the groundwork for more profitable projects down the line. And no matter whether colonization of space is a pipe dream or the manifest destiny of our species, we should not be so stingy as to play science and human exploration against each other. Let each prove its merits. The special report on "The Future of Space Exploration," beginning on page 60, offers suggestions about what may be some of the worthiest goals to set next.

What we are starting to do in space is still not exactly a race, but it is also no longer just a grab for territory and bragging rights, either. It is more like the very earliest stages in settling a frontier: claiming turf but then also extending the reach of society, commerce and personal interests to it, however tentatively. It can be a gritty, unglamorous project. Romantic visions of settling the old West have little relevance; the space frontier will be like nothing humanity has ever experienced. And to the extent that it is a race, it is one we all can win. ■

MATT COLLINS

© 2007 SCIENTIFIC AMERICAN, INC.

THE EDITORS' BLOG

For more posts from SciAm Observations, see *www.SciAm.com/blog*

Are Taxpayers Going to Foot the Bill for Disasters Brought about by Climate Change and Runaway Development?

Posted by Christopher Mims, August 2, 2007

That's a trick question—they already are. As Victoria Schlesinger and Meredith Knight reported in a just-posted exposé for *Scientific American*—"Insurers Claim Global Warming Makes Some Regions Too Hot to Handle"—insurers are dumping coverage of those who may be in the path of global-warming-supercharged storms and rising sea levels.

While the climate change angle of this story is relatively new—insurers will not yet disclose to what degree climate change factors into their calculations when they're deciding to increase prices or in some cases dump coverage of entire areas—what's not new is the accelerating pace of coastal development that has put so much property in harm's way. When big storms like Katrina and Rita come, it's a one-two punch, and the insurers, who are after all for-profit entities, were shaken by the huge losses racked up by these disasters.

Since 1968 the federal government has stepped in to cover individuals who could not get coverage otherwise.

Since its inception, the National Flood Insurance Program (NFIP), which has typically run at a loss, has become the country's primary provider of flood insurance. For instance, in 2005 and 2006 NFIP requested and was granted $24 billion in loans from the U.S. treasury to reimburse Gulf Coast customers for losses caused by Hurricane Katrina. Evan Mills, an environmental and energy systems scientist at Lawrence Berkeley National Laboratory, says it is unlikely NFIP will ever be able to pay back the loan, given that it pulls in an average of only about $2 billion a year in premiums from consumers.

Whereas some researchers have already completed analyses of the current and ultimate cost of global climate change—Sir Nicholas Stern, former chief economist at the World Bank, commissioned a report that puts the figure at $9 trillion—others have countered with what they believe will be the even

➥ *Continued on page 44*

Sustainable Developments

Ending Malaria Deaths in Africa

One of the world's worst killers can be stopped soon if we make the investment

BY JEFFREY D. SACHS

For Africa, the epicenter of the world's malaria scourge, a historic breakthrough in health and economic development is now within reach. A combination of new technologies, new methods of disease control and rising public awareness is poised to bring malaria deaths down by 90 percent or more—if we follow through.

Efforts at malaria control in the 1950s and 1960s successfully used the insecticide DDT and the medicine chloroquine to eliminate the disease in many temperate and subtropical regions. But malaria persisted in the tropics and especially in Africa, where the intensity of transmission is the world's highest for ecological reasons. Africa pays a fearful price for its ongoing malaria burden, not only in more than one million deaths every year but also in significantly reduced economic growth.

Malaria control is the bargain of the planet.

Until very recently, things were getting worse, not better. The malaria parasite became widely resistant to chloroquine. Confusion over DDT's prudent antimalaria application (sprayed as a thin film on the inside walls of houses) and its function as an insecticide in open fields (which is environmentally unsafe and promotes resistance) also curtailed use of the chemical.

The most promising long-term solution is a vaccine, and exciting candidate vaccines are now in clinical trials. Yet even as we await a vaccine, a confluence of advances gives a chance for a breakthrough in the near term. The first is the invention of long-lasting insecticide-treated bed nets, which protect sleeping individuals against indoor nighttime biting. These nets last for five years, unlike earlier nets that needed retreatment every few months.

The second advance, which can save countless lives, is a new generation of highly effective medicines based on artemisinin, an herbal extract discovered by Chinese scientists. (Artemisinin should be used only in combination with more traditional drugs, however, to prevent the onset of resistance in parasites.)

The third advance is a new approach to disease control. In the past, the U.S. government and other donors favored the sale of bed nets at a discount. The result was a very slow uptake of the nets because most African rural households were too poor to buy them. Moreover, the discounts were targeted only for young children and pregnant mothers, the groups most likely to die from malaria. That targeting policy neglected a crucial point: unprotected individuals serve as reservoirs for malaria infection, not only becoming sick themselves but facilitating transmission back to the "protected" groups because the nets are not 100 percent effective.

The new strategy is based on mass free distribution of nets, with one net for every sleeping site. Everybody is protected from illness, and no group is left as a reservoir for transmission. The artemisinin-based medicines should also be available for free within the villages. This approach is highly affordable for donor

BRUCE GILBERT/EARTH INSTITUTE

© 2007 SCIENTIFIC AMERICAN, INC.

THE EDITORS' BLOG

For more posts from SciAm Observations, see *www.SciAm.com/blog*

more enormous cost of changing our ways in order to avert these harms: Lombard Street Research, a for-profit macroeconomic research think tank that advises businesses, has put the figure as high as $18 trillion.

No matter what we do, it's clear that climate change could be a significant drag on the world economy for centuries to come—not to mention the indirect effects such as wars caused by climate change: current, projected and historical.

Run to Starbucks, Get Less Cancer?

Posted by JR Minkel, July 30, 2007

At first glance, this new droplet of research linking caffeine mixed with exercise to protection against skin cancer in mice seems like grounds for excitement. Mice who were fed the equivalent of one to four cups' worth of caffeine per day and also ran on their wheels showed nearly four times as much destruction of ultraviolet-damaged skin cells as sedentary, decaffeinated mice after two weeks. (The paper was published online July 30 in the *Proceedings of the National Academy of Sciences USA*.) Interestingly, the (premium) blend of coffee and exercise conferred more protection than the added effects of caffeine or exercise alone. The two things seem to be feeding on each other in some unknown way.

I can already imagine the marketing campaign: "SPF 15, now with caffeine!" Then people will start making their own by squirting sunscreen into the coffee grinder. It could all get really gross.

Of course, the normal caveats apply here: these are rodents, and we don't know how well this finding will translate to us nor how much cancer protection a given amount of cell self-destruction might confer.

And after a moment's percolation, I conclude that even if the effect holds for people, it is likely to be more of a comfort to coffee drinkers than a cancer cure.

True, coffee drinking has yet to max out: 37 percent of 18- to 24-year-olds drink it, averaging 3.1 cups per day, according to National Coffee Association numbers. But at

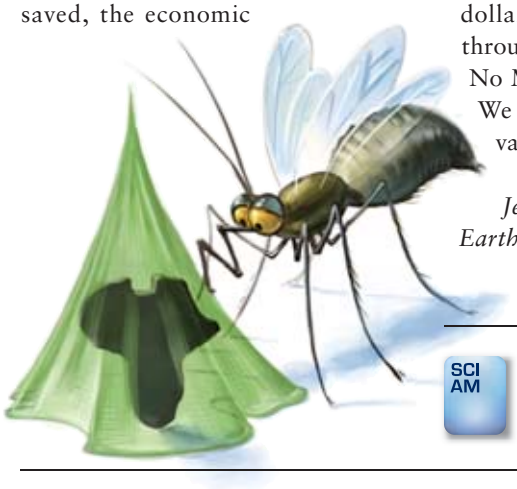
➥ *Continued on page 46*

countries, because the cost of each net is only $5, and each treatment dose of medicine about $1. Gratis distribution of nets is already being applied successfully in several impoverished countries.

Malaria control is the bargain of the planet. A study that my colleagues and I undertook recently showed that comprehensive coverage of nets and medicines, as well as indoor insecticide where advisable, can be accomplished for $3 billion a year in the next few years, which equals just $3 from each person in the high-income world. And these costs will come down in later years as infection rates decline. In addition to the lives saved, the economic gains in Africa would soon amount to tens of billions of dollars a year, manifested in direct reductions of the cost of illness and increased economic growth.

Funding sources are coming into line. The Global Fund to Fight AIDS, Tuberculosis and Malaria is a natural funder and leader. The World Bank can play a pivotal role, especially because the bank's new president, Robert B. Zoellick, has shown leadership on this issue in the past. The Bush administration has recently increased malaria funding. The private sector is ready to step up with support in various ways, and the public is already donating tens of millions of dollars to buy bed nets for the poor through organizations such as Malaria No More (www.malarianomore.org). We are at the threshold of a great advance. It is now time to cross it. ■

Jeffrey D. Sachs is director of the Earth Institute at Columbia University (www.earth.columbia.edu).



MATT COLLINS

SCI AM An expanded version of this essay is available at **www.SciAm.com/ontheweb**

Skeptic

The Really Hard Science

To be of true service to humanity, science must be an exquisite blend of data, theory and narrative

BY MICHAEL SHERMER

BRAD SWONETZ

Over the past three decades I have noted two disturbing tendencies in both science and society: first, to rank the sciences from "hard" (physical sciences) to "medium" (biological sciences) to "soft" (social sciences); second, to divide science writing into two forms, technical and popular. And, as such rankings and divisions are wont to do, they include an assessment of worth, with the hard sciences and technical writing respected the most, and the soft sciences and popular writing esteemed the least. Both these prejudices are so far off the mark that they are not even wrong.

I have always thought that if there must be a rank order (which there mustn't), the current one is precisely reversed. The physical sciences are hard, in the sense that calculating differential equations is difficult, for example. The variables within the

© 2007 SCIENTIFIC AMERICAN, INC.

THE EDITORS' BLOG

For more posts from SciAm Observations, see *www.SciAm.com/blog*

this stage, few of us need to be reminded that exercise is good for you, and coffee already has a mixed bag of potential health benefits (possible reduced risk of diabetes) and detriments (hypertension) that seems unlikely to sway anybody one way or the other, assuming they weigh the evidence.

Even one of the researchers who identified the link between coffee and diabetes didn't advocate drinking more java, according to the kicker in a *New York Times* piece from last year.

Maybe pharma could identify some caffeine-ish compound that conferred extra protection against skin cancer for those with a family history of it. But until then, keep using your (caffeine-free) sunscreen, kiddies.

Want to Get from Point A to Point B in One Piece? Don't Take the Shuttle

Posted by Christopher Mims, July 20, 2007

As a follow-up to my previous post about the likelihood of being killed by various forms of transport, I looked up statistics on the space shuttle. Here's how it breaks down:

Fatalities per 100 million passenger-kilometers:

Car: 1.1

Rail: less than 0.1

Air: 0.1

Space shuttle: about 1.9

That assumes that all shuttles have cumulatively logged approximately 718 million kilometers in their many orbits around Earth and that there have been two lost missions, for a total of 14 fatalities.

Fatalities per 100 million passenger-hours:

Car: 32

Rail: 2

Air: about 35

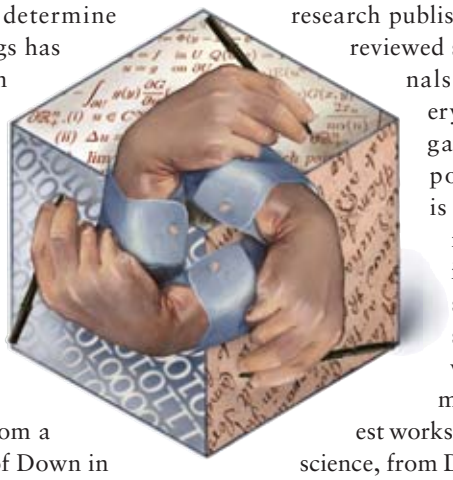
Space shuttle: 52,599

That assumes that the shuttle has logged 11,000 days in flight.

Anyway, there you have it, folks: you're 1,600 times more likely to die on a two-day space shuttle trip than on a week's hard driving cross-country (assuming you're clocking eight to 10 hours a day on the road).

causal net of the subject matter, however, are comparatively simple to constrain and test when contrasted with, say, computing the actions of organisms in an ecosystem or predicting the consequences of global climate change. Even the difficulty of constructing comprehensive models in the biological sciences pales in comparison to that of modeling the workings of human brains and societies. By these measures, the social sciences are the hard disciplines, because the subject matter is orders of magnitude more complex and multifaceted.

Between technical and popular science writing is what I call "integrative science," a process that blends data, theory and narrative. Without all three of these metaphorical legs, the seat on which the enterprise of science rests would collapse. Attempts to determine which of the three legs has the greatest value is on par with debating whether π or r^2 is the most important factor in computing the area of a circle.



Consider data and theory first. I began this column in April 2001 with what I called "Darwin's dictum," which came from a quote from the sage of Down in response to a critique that *On the Origin of Species* was too theoretical and that he should have just "put his facts before us and let them rest." Darwin responded by explaining the proper relation between data and theory: "About thirty years ago there was much talk that geologists ought only to observe and not theorize, and I well remember someone saying that at this rate a man might as well go into a gravel-pit and count the pebbles and describe the colours. How odd it is that anyone should not see that all observation must be for or against some view if it is to be of any service!"

Charles Darwin's dictum holds that if observations are to be of any use they must be tested against some view—a thesis, model, hypothesis, theory or paradigm. The facts that we measure or perceive never just speak for themselves but must be interpreted through the colored lenses of ideas. Percepts need concepts, and vice versa. We can no more separate our theories and concepts from our data and percepts than we can find a true Archimedean point—a god's-eye view—of ourselves and our world.

Data and theory are not enough. As primates, humans seek patterns and establish concepts to understand the world around us, and then we describe it. We are storytellers. If you cannot tell a good story about your data and theory—that is, if you cannot explain your observations, what view they are for or against and what service your efforts provide—then your science is incomplete. The view of science as primary research published in the peer-reviewed sections of journals only, with everything else relegated to "mere popularization," is breathtakingly narrow and naive. Were this restricted view of science true, it would obviate many of the greatest works in the history of science, from Darwin's *On the Origin of Species* to Jared Diamond's *Guns, Germs, and Steel,* the evolutionary biologist's environmental theory about the differential rates of development of civilizations around the world for the past 13,000 years.

Well-crafted narratives by such researchers as Richard Dawkins, Steven Pinker, the late Stephen Jay Gould and many others are higher-order works of science that synthesize and coalesce primary sources into a unifying whole toward the purpose of testing a general theory or answering a grand question. Integrative science is hard science. ■

Michael Shermer is publisher of Skeptic *(www.skeptic.com). His latest book is* Why Darwin Matters.

MATT COLLINS

© 2007 SCIENTIFIC AMERICAN, INC.

Anti Gravity

Carrots, Sticks and Robot Picks

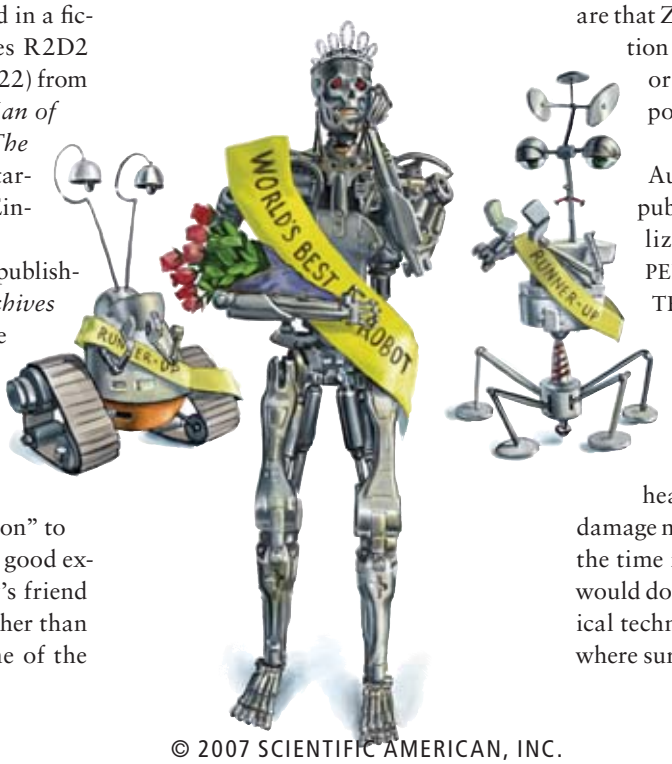
Some strange science stories of recent vintage

BY STEVE MIRSKY

Ahhh, midsummer, when space travelers fail their sobriety tests. Of course, NASA's inebriated astronauts debacle was well covered. Here are some perhaps lesser-known tales of whoa. For instance, in late July, the *Times of London* published its list of the 50 best movie robots ever, in conjunction with the release of the movie *Transformers*. Little did I know that I owned a transformer—I left the car lights on all night *twice* last week, and my vehicle turned into a really big paperweight. Anyway, the *Times* picks the Terminator as its best robot, with the HAL 9000 coming second (an insane computer is not really a robot, to my human mind) and KITT, the talking car from *Knight Rider,* finishing third (again, not really a robot, although having William Daniels's voice remind me to shut the lights off would really come in handy). These choices can only be considered absurd in a fiction-filled universe that includes R2D2 (the *Times*'s #11), the Fembots (#22) from *Austin Powers: International Man of Mystery* and Gort (#13) from *The Day the Earth Stood Still,* also starring Sam Jaffe's hair as Albert Einstein's hair.

Speaking of lists, researchers publishing in the August issue of the *Archives of Sexual Behavior* delineated the results of their survey of more than 2,000 people and announced their exhaustive compilation of the 237 reasons that people have sex. Justifications ranged from "to show my affection" to "it feels good" to "it seemed like good exercise." That's right, somebody's friend with benefits ranks just a bit higher than an elliptical trainer. Oddly, one of the most famous reasons in history fails to make the list: "I was fulfilling prophecy, having already killed my father and married my mother." As Homer (Simpson) famously asked of the Oedipus account, "Who pays for that wedding?"

Speaking of parents and problem children, here's an excellent experiment to perform on any three-year-old whose parents are constantly telling you how smart the kid is. Take a food item—a couple of carrots, for example—and put one in an unmarked bag. Put the other one in a McDonald's bag. Then have the little genius taste both and ask which carrot was better. Or save yourself all this trouble by reading the August issue of *Archives of Pediatrics & Adolescent Medicine,* in which researchers found that the $10 billion dished out every year in the U.S. by food and beverage companies to market to small children is money well spent. Because 54 percent of preschool kids surveyed preferred the alleged McDonald's carrot, whereas only 23 percent liked the carrot in plain wrapping better. The effect was magnified when the test food was french fries: 77 percent said McDonald's-looking potatoes, only 13 percent said the other potatoes, and 10 percent said let's call the whole thing off.



Well, at least the kids aren't chewing on Developmentally Delayed Elmo, Intensive Care Bears, the Big Bag o' Paint Chips or any of the millions of toys recalled over the summer because they contained unsafe levels of lead. The toys were manufactured in China, which had already endeared itself to international consumers with recent exports of tainted toothpaste, contaminated pet food, phony and dangerous medications, and enough other shoddy products to cause the actual state execution of the director of that country's Food and Drug Administration. The method was not revealed, but chances are that Zheng Xiaoyu died from an injection of some seriously unsafe drugs or from exceptionally rapid lead poisoning.

In a distantly related story, on August 6 the Reuters wire service published an article with the tantalizing headline "GERMAN HAS PENCIL IN HEAD REMOVED AFTER 55 YEARS." According to the report, Margret Wegner fell while carrying the pencil when she was four years old. "The pencil went right through my skin—and disappeared into my head," she remembered. With the damage miraculously minimal, doctors at the time feared that getting the lead out would do more harm than good. But medical technology finally reached the point where surgeons could reach the point. ■

PHOTOGRAPH BY FLYNN LARSEN; ILLUSTRATION BY MATT COLLINS

© 2007 SCIENTIFIC AMERICAN, INC.

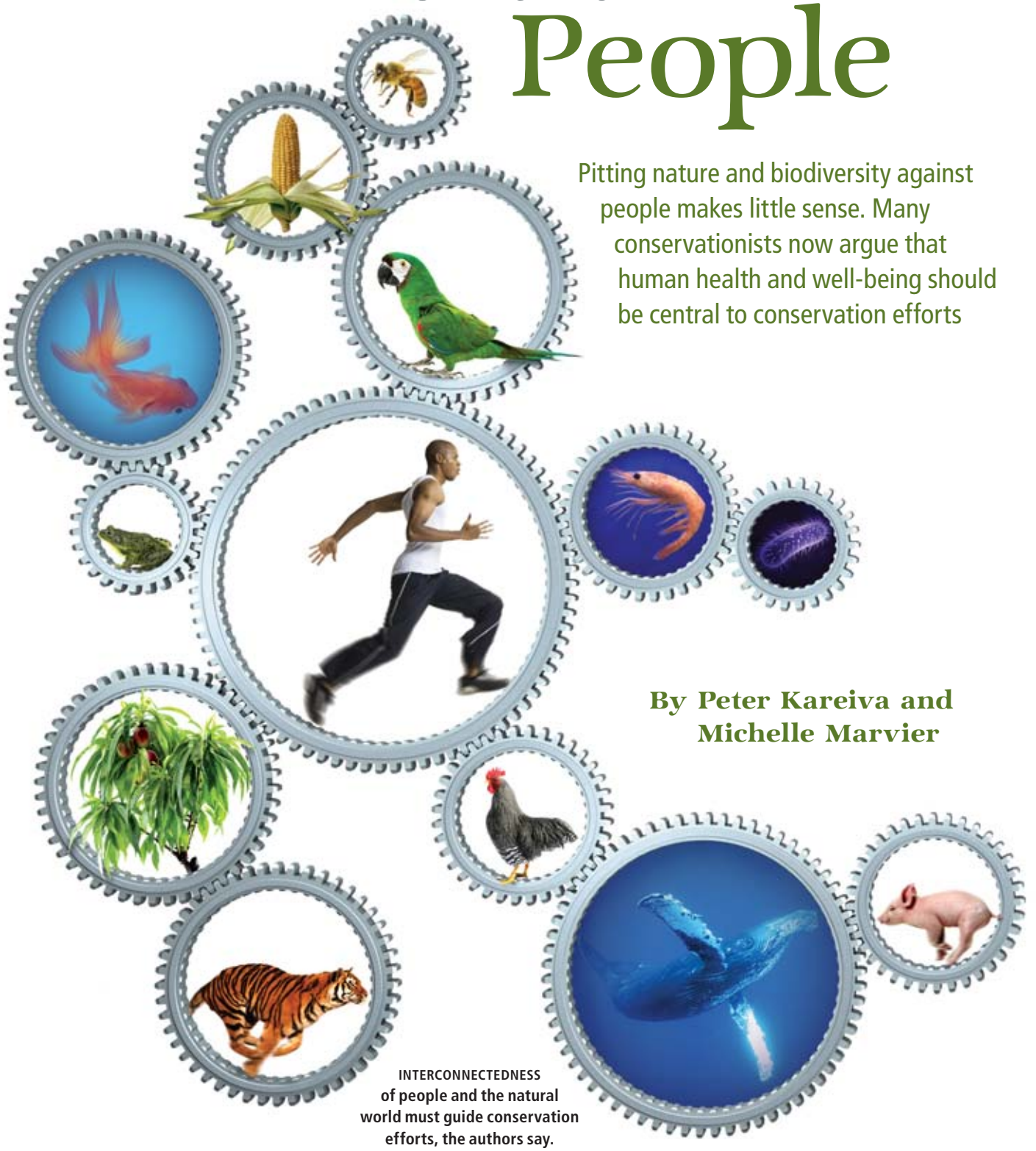
Page Intentionally Blank

SCIENTIFIC AMERICAN Digital

Conservation for the People

Pitting nature and biodiversity against people makes little sense. Many conservationists now argue that human health and well-being should be central to conservation efforts

By Peter Kareiva and Michelle Marvier



INTERCONNECTEDNESS of people and the natural world must guide conservation efforts, the authors say.

© 2007 SCIENTIFIC AMERICAN, INC.

KOMODO NATIONAL PARK in Indonesia has local support because it led to income from cultivating reef fish and selling carvings.

In 2004 the World Conservation Union placed three vultures—the long-billed, the slender-billed and the Oriental white-backed—on the critically endangered list. Populations of all three reached nearly 40 million in India and South Asia in the early 1990s but had fallen by more than 97 percent. The reasons for saving these vultures from extinction could be framed in familiar terms: we have an ethical obligation to save the world's biodiversity for its own sake. But the reasons could also be outlined in a less familiar way.

For a long time, observers did not know what was causing the vultures' decline. Some speculated the culprit was habitat loss or pollution. Several years ago researchers discovered that the birds were being killed by an anti-inflammatory drug, diclofenac, commonly administered to cows. In bovines and humans, the medicine reduces pain; in vultures, it causes renal failure. As the vultures have disappeared, hundreds of thousands of cow carcasses customarily left for the birds have festered in the sun, where they incubate anthrax, according to some reports, and are consumed by dogs. Because of the ready food supply, the feral dog population has exploded—and with it the threat of rabies. Thus, the vultures' fate may be linked with that of millions of people; saving the vultures from extinction would protect people from dangerous disease.

Casual observers do not always see links between human well-being and aiding endangered species, but such connections abound in many situations that engage conservationists. Ecosystems such as wetlands and mangrove stands protect people from lethal storms; forests and coral reefs provide food and income; damage to one ecosystem can harm another half a world away as well as the individuals who rely on it for resources or tourism revenue.

Despite these mutual dependencies, the public and some governments increasingly view efforts to preserve biological diversity as elevating the needs of plants and animals above those of humans. To reverse this trend—and to better serve humanity and threatened organisms—we and a growing number of conservationists argue that old ways of prioritizing conservation activities should be largely scrapped in favor of an approach that emphasizes saving ecosystems that have value to people. Our plan should save many species, while protecting human health and livelihoods.

Out, Out Hot Spot

Conservation's misanthropic reputation has arisen, in part, because millions of people have been forced off their land or have otherwise had their sources of food and income snatched from them so that animals and habitats could be preserved. Kenya's president Mwai Kibaki's con-

Komodo dragon

KEY CONCEPTS

- Preserving biodiversity for its own sake, particularly in areas called hot spots, is not working as a conservation strategy.
- Focusing on protecting ecosystems vital to people's health and material needs makes more sense.
- Such ecosystems would include not only forests but also wetlands that maintain clean water, mangroves that shield against storms and reefs that sustain fisheries.
- Saving these sites can preserve biodiversity and ensure that people are a priority. *—The Editors*

JURGEN FREUND *Nature Picture Library (reef)*; JEZ O'HARE (*craft workers*); JOE McDONALD *Corbis (dragon)*

© 2007 SCIENTIFIC AMERICAN, INC.

[THE BASICS]

A TALE OF TWO STRATEGIES

The well-known conservation strategy of saving hot spots (*left*) does not protect many ecosystems valuable to human health and development. An ecosystem services approach (*right*) would do just that and would establish a different metric for setting priorities.

HOT-SPOTS STRATEGY

THE BASIC IDEA

Identify threatened areas with high plant diversity and protect them, under the assumption that doing so protects an array of animals—which are often harder to catalogue than plants. To date, 25 such hot-spot regions have been named, including Bocaina National Park in Brazil (*above*).

TYPICAL APPROACH

Establish a national park or reserve to protect animal and plant life. Discourage people from living on or using that land. Patrol and enforce boundaries.

DRAWBACKS

Areas rich in plant species are not necessarily rich in animal diversity. Local people are often displaced or lose important resources. Hot spots have not captured the public's imagination or support.

ECOSYSTEM SERVICES STRATEGY

THE BASIC IDEA

Make clear people's dependence on various ecosystems—as is the case with tourism revenues in Punta Tomba, Argentina (*above*)—and identify ecosystems that are gravely threatened and whose impairment will harm local residents.

TYPICAL APPROACH

Where ecosystems are being degraded, establish a conservation plan that protects the ecosystem and benefits the community dependent on it.

WHY IT'S A BETTER IDEA

As people see more clearly their reliance on various ecosystems for their health and economic security, they will support conservation projects. As a result, biodiversity will be preserved, but not at the expense of humans.

WHAT ARE ECOSYSTEM SERVICES?

In a recent global study, the United Nations identified four kinds of services:

PROVISIONING—supplying food or genetic resources, for example

REGULATING—providing flood control, climate modulation or other similar functions

CULTURAL—offering benefits that are nonmaterial, such as a sense of place and spiritual well-being

SUPPORTING—delivering the most basic elements of an ecosystem, including nutrient cycling, soil formation or pollination

troversial decision to return Amboseli National Park to its original Maasai inhabitants reflects a growing discontent with such preemptory displacements. It is a global discontent. Hunters and farmers in Asia and throughout Africa contend that parks limit their diet and earnings. U.S. farmers and loggers are angry about losing their water privileges or their jobs because of salmon and spotted owls.

Public perception of a nature-versus-people theme also derives from the conservation strategy of focusing on so-called hot spots. In 1988 Norman Myers of the University of Oxford developed the idea of biodiversity hot spots, small areas that harbor a great variety of endemic, or native and geographically restricted, plant species. Myers used diversity of such plants as the measure because plant lists were the most reliable and often the only data available and because it was thought that plant diversity served as a good proxy for animal diversity. Myers and his colleagues at Conservation International went on to identify 25 hot spots—the Brazilian Cerrado region is one; the Horn of Africa is another—on which to focus conservation projects.

Earlier conservation campaigns had centered on charismatic species such as pandas, whales and seals. In contrast, the concept of hot spots provided a set of rigorous, quantifiable criteria by which to guide conservation investment—a triage system based on counting species was more scientific than one based on compelling photographs of cute or iconic animals. The approach seemed more realistic as well: conservation organizations have limited funds and could now put money in places where the most species might be saved. For the past 15 years, this strategy has been embraced by philanthropic and multinational organizations alike.

Although "hot spot" is a compelling phrase, the idea of biodiversity that underlies it has not succeeded in capturing public imagination or interest. One recent survey showed that only 30

TUI DE ROY *Minden Pictures* (*left*); GITAN JEAN (*right*)

© 2007 SCIENTIFIC AMERICAN, INC.

percent of Americans have heard of the term "biodiversity." And many people working in conservation are careful to avoid the word because it engenders apathy or a negative response. Biodiversity hot spots clearly are not galvanizing the public to fund or participate in conservation.

Some scientists are not so keen on hot spots either. C. David L. Orme of Imperial College London recently pointed out that they might be false advertising: places with a lot of native plant species do not necessarily have many butterfly or vertebrate species as well. Marcel Cardillo, also at Imperial College London, has noted that animals in floral hot spots are not those most vulnerable to extinction: that distinction goes instead to mammals in boreal forests and arctic regions.

Other biologists have shown that many of the world's least diverse regions provide important seasonal homes, migratory stops or nesting sites. Half a million Magellanic penguins gather each September in Punta Tomba, Argentina, for instance. This dry, shrubby region of Patagonia is home to few endemic plants and could not be called even a biodiversity lukewarm spot. Yet the penguins nesting there are critical to the local economy; 70,000 tourists visit every year to see them. Many similar places exist, sites of low plant biodiversity that are nonetheless crucial to species with far-reaching ecological or economic importance: stretches of tundra that support ducks, swans and geese; temperate rivers where salmon spawn.

A Service Paradigm

Conservation needs additional principles to guide it. Although people may not comprehend the concept of biodiversity, they do value nature as a source of food, fuel, building materials, recreation and inspiration. Ecologists have begun to quantify this natural capital under the umbrella of "ecosystem services," a term coined by Paul R. Ehrlich and championed by Gretchen C. Daily, both at Stanford University. These services include products for which there are markets, such as medicines and timber, as well as processes whose economic value usually goes unconsidered: water filtration, pollination, climate regulation, flood and disease control, and soil formation. When Robert Costanza of the University of Vermont and other economists attempted to place a dollar value on those processes, they found that the yearly value of such economic services outstripped the gross domestic product of all countries combined.

COUNTRIES highlighted here harbor life raft ecosystems that are conservation priorities for the authors. Such ecosystems are ones whose conservation and restoration would dramatically improve people's lives. The authors identified the sites using data on poverty, the importance of natural resources to the economy, and the extent of land degradation.

The idea of focusing on ecosystem services is not being embraced by academics alone; increasingly, governments and nongovernmental organizations are considering protecting these services a fundamental goal. In 2000 the United Nations called for a study of ecosystem services. A year later an international team of more than 1,300 scientists undertook one of ecology's most ambitious endeavors: the Millennium Ecosystem Assessment. The project documented the impacts humans have had on ecosystem services in the past 50 years. Services were divided into four categories: provisioning (supplying products such as food or genetic resources), regulating (contributing regulatory functions such as flood control), cultural (supplying nonmaterial benefits such as a sense of spiritual well-being) and supporting (providing basic elements of the ecosystem, such as soil formation). The assessment found that most ecosystem services not only have declined but are being used unsustainably.

For the wider public, the Indian Ocean tsunami in 2004 and Hurricane Katrina in 2005 brought into sharp focus the relation between ecosystems and human living conditions. In both cases, damage was amplified by loss of natural vegetation. The destruction over the past 70 years of some 1,500 square miles of Louisiana's marshlands and eelgrass beds greatly exacerbated the storm surge generated by Katrina [see "Drowning New Orleans," by Mark Fischetti; SCIENTIFIC AMERICAN, October 2001]. In Southeast Asia the widespread conversion of coastal mangrove forests into shrimp ponds meant there was no

WHEN INDIA'S VULTURES started dying, the ramifications for humans were not immediately clear—until the threat of rabies from increased numbers of feral dogs feasting on animal carcasses made the link explicit. Saving wildlife often saves people, too. Connections such as this between the fate of wildlife and the health of humans abound.

PAGE 50: SAMUEL AND PEDRO VELASCO *SW Infographics*; DAVIES & STARR *Getty Images (corn)*; DON FARRALL *Getty Images (bee and goldfish)*; GEORGE DOYLE AND CIARAN GRIFFIN *Getty Images (parrot)*; RUBBERBALL PRODUCTIONS/GETTY IMAGES *(man)*; PURESTOCK/GETTY IMAGES *(shrimp, rooster and frog)*; MEDICALRF.COM/GETTY IMAGES *(microorganism)*; FLIP NICKLIN *Minden Pictures/Getty Images (whale)*; DIGITAL VISION/GETTY IMAGES *(piglet)*; JOSEPH VAN OS *Getty Images (tiger)*; CREATIV STUDIO HEINEMANN *Getty Images (plant)*; THIS PAGE: ANN SANDERSON *(top)*; ASHOK JAIN *Nature Picture Library (vultures)*

© 2007 SCIENTIFIC AMERICAN, INC.

[CASE STUDY]

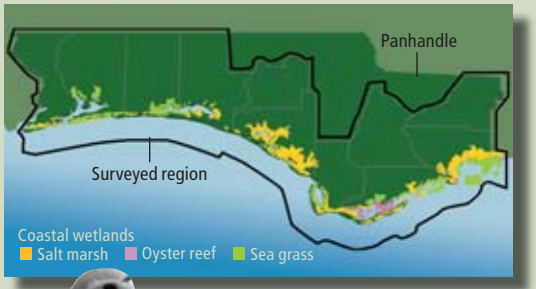
PROTECTING POOR COMMUNITIES AND HABITATS

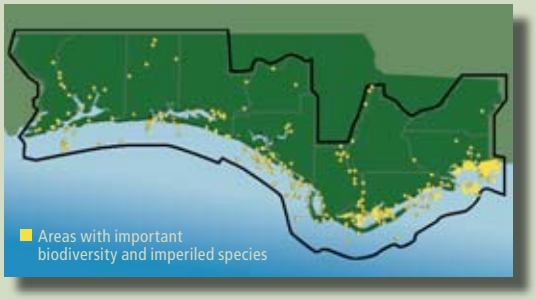
THE PROBLEM

The salt marshes, sea-grass beds and oyster reefs of Florida's Gulf Coast harbor manatees, sea turtles, piping plovers and many other threatened species, as well as serving as nurseries for economically important shrimp, crab and red snapper. These habitats also provide protection from storm surges that accompany hurricanes, such as Dennis in 2004 (*photograph at right*). Yet strategies to defend and restore coastal ecosystems—which could simultaneously assist people and expand habitats for threatened and economically valuable species—have largely been ignored in favor of engineering projects that accelerate erosion and habitat loss.

THE SOLUTION

Scientists from the Nature Conservancy and the National Oceanic and Atmospheric Administration recently combined maps of critical habitats and threatened species in the Florida Panhandle with maps of anticipated storm surges and of communities most likely to suffer because of storms (*below*). By overlaying these data sets, they were able to identify areas whose restoration should simultaneously protect the most vulnerable human populations as well as many of the areas' most important species.





Piping plover

wave buffer to protect against the tsunami. Post-tsunami studies led by Sri Lankan researcher Farid Dahdouh-Guebas of Vrije University in Brussels found that shorelines with intact mangrove forests suffered almost no damage. Neither Louisiana's marshes nor Sri Lanka's mangroves rank among the world's biodiversity hot spots: they have virtually no endemic plant species, and we estimate that the number of plant and animal species they contain does not approach one tenth of that found in a rain forest.

Connections between habitat loss and economic loss that are not always as obvious can also be significant. The winds that whip through Africa's ever expanding Sahara and Sahel carry dust that is blown west over the Atlantic Ocean. Every year several hundred million tons of such sand land in the Americas or the Caribbean. Once there the dust, pollutants, microorganisms and nutrients accompanying the sand play a part in wiping out coral reefs—reducing tourism and fisheries. Overgrazing and unsustainable farming practices in northern and sub-Saharan Africa have fueled poverty, famine and malnutrition regionally and undermined corals and economies half a world away.

The economic benefits afforded by ecosystem services are most needed by developing nations. These countries derive substantial income from timber, fiber and agriculture; forestry and fisheries are typically five to 10 times more important as components of national economies for such nations than for the U.S. and Europe. A 2005 U.N. report convincingly explained that maintaining the environment is the key to alleviating poverty for the world's 750 million rural poor.

Human health is also threatened when ecosystems and natural cycles break down. The vultures of India are but one example among hundreds. Almost two million people die every year because of inadequate or unclean water supplies. Conserving wetlands and forests would reduce these deaths: wetlands provide natural filters that improve water quality for drinking and agriculture; healthy forests lock up sediment that would otherwise muddy water. Saving forests and grasslands would reduce plumes of dust originating in Africa and the even larger ones crossing the Pacific Ocean from western China that recently have been linked to a rise in U.S. asthma cases.

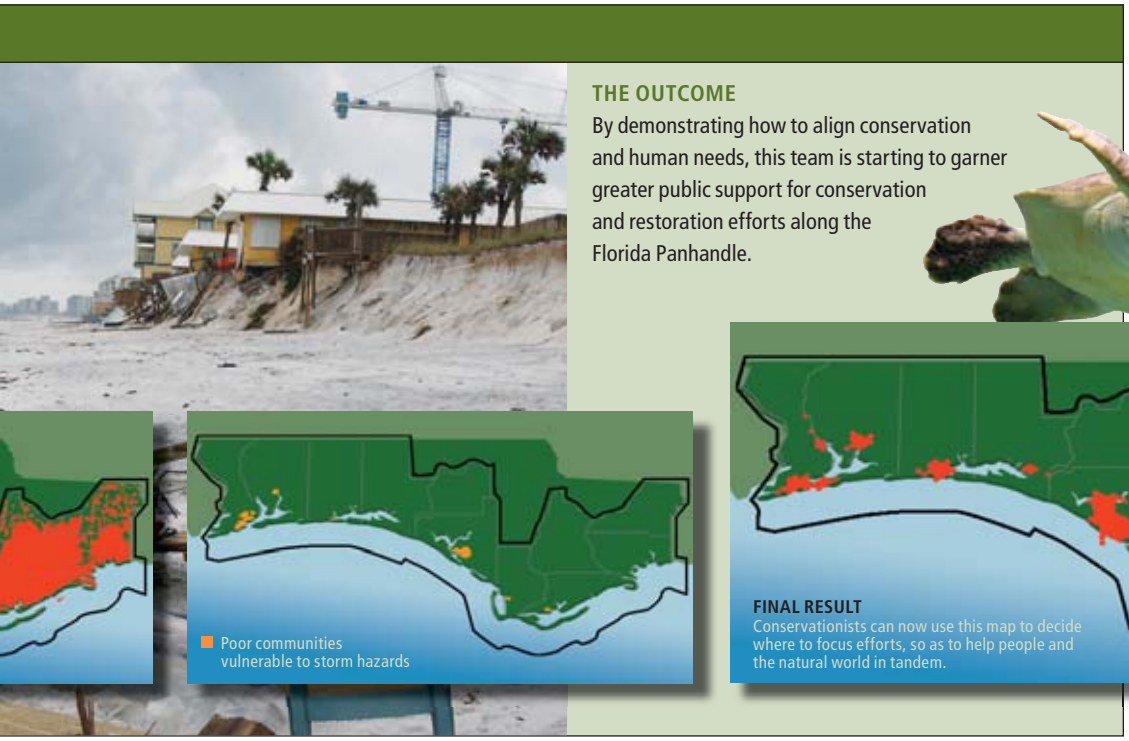
A subtler connection between ecosystem degradation and human health can be seen in disease-causing organisms that move from wildlife to humans. Two thirds of the world's emerging diseases, such as the Ebola virus and avian flu, are caused by pathogens that infect nonhuman animal hosts and only make contact with people because of changes in land use and agricultural practices. At issue are not just "exotic" diseases, however. By eliminating wolves and

[THE AUTHORS]

Peter Kareiva and Michelle Marvier have worked together for many years, conducting studies of transgenic crops and of salmon in the Pacific Northwest. They are now collaborating on a conservation textbook. Kareiva is chief scientist at the Nature Conservancy, where he conducts research and travels widely to teach and advise on international projects. Marvier is a professor at Santa Clara University, where she directs the Environmental Studies Institute. Both feel strongly that conservation must be more directly connected to people.

ANN SANDERSON (*maps*); JIM BRANDENBURG *Minden Pictures* (*piping plover*); MARI DARR WELCH *AP Photo* (*hurricane*); COURTESY OF PETER KAREIVA (*Kareiva and Marvier*)

© 2007 SCIENTIFIC AMERICAN, INC.



mountain lions, people in the eastern U.S. triggered an explosion in the deer and deer tick populations, which has resulted in more than 20,000 new cases of Lyme disease annually. Attempts to eradicate predators more than a century ago have jeopardized human health today.

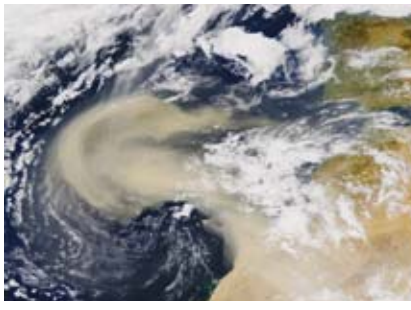
Life Raft Conservation

The focus on ecosystem services we are advocating is, in many ways, a repackaging of traditional conservation ideas that emphasize interconnectedness. But our approach differs in several significant ways. First, we believe many conservationists are in denial about the state of the world and must stop clinging to a vision of pristine wilderness. One quarter of a million people join the planet every day. More forests and wetlands will be cleared for agriculture, and more ocean species will be fished to depletion. Biodiversity is going to decline. Wilderness separate from human influence no longer exists.

Because our environment will consist mainly of human-influenced systems, biodiversity protection must be pursued in the context of landscapes that include urban centers, intensive agriculture, and managed forests and rivers, not just nature preserves. Ironically, protected areas will most likely need to be intensely supervised to retain their "wildness." Managers in many parks have come to recognize this reality. Kruger National Park in South Africa is a highly managed landscape where natural watering holes have been replaced by drilled wells and elephant populations are culled to prevent overcrowding.

The second major shift we urge is that conservationists focus foremost on regions where the degradation of ecosystem services most severely threatens the well-being of people: stands of mangroves in Asia, marshes in the southeastern U.S., drylands in sub-Saharan Africa and coral reefs around the world. This approach would be especially expeditious where government agencies and conservation groups seek to work together for both public protection and for conservation. For example, in the Florida panhandle, a partnership between the National Oceanic and Atmospheric Administration and the Nature Conservancy is working to identify areas of joint concern for public well-being and traditional conservation. By mapping habitats in terms of their ability to protect human communities in addition to their biodiversity, participants are finding important areas to preserve.

Third, conservationists should collaborate more closely with development experts. In the past two decades, many sustainable development projects have sought to bring these groups' interests together, but only with attention to already marketed items, such as fish or nontimber forest products—and rarely with the spectrum of ecosystem services in mind. By combining and coordinating the energy and capital of



DUST from degraded grassland ecosystems in sub-Saharan Africa travels far afield in wind, harming coral reefs, tourism and fisheries in the Caribbean. Protecting important ecosystems in one part of the world can also help people an ocean away.

J. AND A. KOSTEN/Peter Arnold, Inc. (*turtle*); NASA (*dust*)

© 2007 SCIENTIFIC AMERICAN, INC.

conservation forces and human welfare projects, the experts could enhance the efficiency and impact of both efforts. For instance, investments in clean, sediment-free water are often the same investments capable of protecting aquatic biodiversity.

Without a close connection between conservation and social issues, policies that protect biodiversity are unlikely to find much public support. Michael Shellenberger and Ted Nordhaus of the consulting company American Environics made the point in a 2004 essay entitled "The Death of Environmentalism" that environmental groups need to move beyond their tendency to put the environment in an airtight container away from the concerns of others—or be doomed to irrelevance. We believe this stinging indictment of environmentalism applies equally to the conservation movement.

Finally, the conservation efforts we envision will be assessed not just by the number of species protected but by improvements to people's well-being. Such assessments are already beginning. The Indonesian government and the Nature Conservancy came together to set aside Komodo National Park in 1980, partly to protect the threatened Komodo dragon and partly to preserve forests and coral reefs. Park admission fees were directed to local development projects and to new sources of income: seaweed cultivation, tourism, wood carving and the breeding of prized reef fish. A 2006 survey of local villagers bordering Komodo found that the overwhelming majority was highly supportive of the protected area because of the new income it had generated.

Lurking Unease

Some people will be alarmed by this proposal because the services provided by nature do not always correlate with biodiversity. A second source of anxiety about our approach is the fact that the plants and animals most central to ecosystem services and human economy tend to be fairly abundant. But rare species still have a crucial role: as insurance. With global climate disruption and massive modification of land, the rare species of today may become the abundant species of tomorrow, and so we should save as many as possible. In California, the nonnative European honeybee is the most important pollinator from an economic perspective. If the European honeybee population were to become dramatically reduced (and it has recently been threatened by introduced mites), some of the less abundant native bees might increase and fill the vital economic role of crop pollinators.

Although it would be morally reprehensible for humans to allow the extinction of all species except those few that provide, or might provide, services, it is also unrealistic to think we can return any substantial part of the world to a preindustrial state. Some human-caused extinctions are inevitable, and we must be realistic about what we can and cannot accomplish. We must be sure to first conserve ecosystems in places where biodiversity delivers services to people in need.

[CASE STUDY]

PROTECTING DRINKING WATER

THE PROBLEM

Much of the water that supplies Quito, Ecuador's largest city, originates in Andean highlands that are home to a tremendous variety of endemic plants and animals, including the spectacular Andean condor. Although a condor reserve has been set up (*photograph*), enforcement is poor. Downstream, many areas around the city do not have enough water to meet their needs, and most of the city's monitored watersheds have undrinkable water. Poor farming and logging practices in the vicinity of the condor reserve and farm animals grazing too close to stream and river channels are the culprits.

THE SOLUTION

In 2000 the U.S. Agency for International Development (USAID), the Nature Conservancy and local Ecuadorian partners established a water fund. Quito's hydropower firm, the Andina beer company, Quito's municipal water supplier and a 2 percent tax on Quito's residents provide the income. The fund has collected $4.9 million for supporting conservation, education and water projects upstream from Quito.

THE OUTCOME

As of this year, 11 new park guards have been hired to patrol the protected area, and a massive education program with farmers has been undertaken to teach better land-use practices. More than 3.5 million trees have been planted in an effort to reforest denuded watersheds. It is too early to know whether the more sustainable practices are yielding the desired improvements in water quality, but a network of hydrological monitoring stations is being created. Public enthusiasm for water conservation has grown dramatically.

PETE OXFORD *Nature Picture Library*

© 2007 SCIENTIFIC AMERICAN, INC.

[CASE STUDY]

PROTECTING WILDLIFE

THE PROBLEM

In Namibia the marginalized San people, often referred to as bushmen, suffer extreme poverty and one of the highest rates of stunting among children. They have been displaced from their communal lands and, left with no sustainable way of making a living, have been forced into poaching and excessive hunting. The black rhino, one of the most endangered species in the world, has been one casualty.

FRANS LEMMENS *Peter Arnold, Inc.*

THE SOLUTION

In 1996 the government of Namibia passed an act giving the indigenous people ownership of game animals and all revenues from tourism and game products. Local conservancies covering 17 percent of the Namibian land and including 60 communities were set up to manage the wildlife, tracking game movements, for example. USAID provided funds to help the San establish and participate in these local conservancies.

THE OUTCOME

Where the local conservancies are active, wildlife is rebounding, with 600 percent increases in populations of elephants, zebras, oryx and springbok. Namibia also now has the world's largest free-roaming black rhino population. At the same time, more than 500 full-time jobs and more than 3,000 part-time jobs have been created for the local people. In 2004 tourism (*left*) and hunting generated $2.5 million in income. This case also illustrates some of the challenges of the ecosystem services approach to conservation: many San remain marginalized, and some observers argue that the indigenous people should be given ownership of the land, not just of the wildlife.

We suggest that instead of mapping the top 10 or 25 locations in need of protection in terms of native plant richness, conservationists should generally seek to identify life raft ecosystems—areas with high rates of poverty, where a large portion of the economy depends on natural systems and where ecosystem services are severely degraded. Conservation efforts aimed at providing clean water, reducing soil erosion and preventing overfishing will help people and protect much, though certainly not all, biological diversity. These types of projects will generate a much broader support base than is typical of most conservation efforts.

Meanwhile some dedicated organizations should continue to support the conservation of species and places without conspicuous utility. Shifting emphasis to ecosystem services does not mean totally changing conservation goals; it means broadening public support for conservation and fostering a shift in emphasis for organizations able to make that shift.

Natural Economy

Whether efforts that conserve ecosystem services are able to support economic development remains to be proved. The future of ecosystem services as a conservation strategy may depend on the unlikely collaboration of ecologists and finance experts. Indeed, much of the enthusiasm for this approach is coming from the business community. In November 2005, for example, the Goldman Sachs Group announced an ecosystem services framework for its own business operations, which included making $1 billion available for investing in renewable energy, assessing the impacts of its projects on ecosystem services as standard operating procedure, and establishing a think tank to explore green markets.

The World Bank is also encouraging nations to embrace green accounting methods in which economic assets and national productivity assessments include measures that credit environmental and ecosystem services and subtract degradation that results from pollution or destructive extraction. Economic valuation and the creation of markets for ecosystem services offer the possibility of providing a quantifiable conservation metric to which corporations and people can readily relate, an improvement over policies based on charismatic species or plant endemism.

A few enlightened voices—such as those of 2004 Nobel Peace Prize winner Wangari Muta Maathai and former U.N. secretary-general Kofi A. Annan—have called attention to the connection between the environment, human prosperity and peace. Annan has stated that "our fight against poverty, inequality and disease is directly linked to the health of the earth itself." Conservationists need to hear and communicate this message. Conservation will only become truly global and widely supported when people are central to its mission. ■

MORE TO EXPLORE

Ecology for a Crowded Planet. M. Palmer et al. in *Science,* Vol. 304, pages 1251–1252; May 28, 2004.

How Much Is an Ecosystem Worth? Assessing the Economic Value of Conservation. World Bank, 2005.

Valuing Ecosystem Services: toward Better Environmental Decision Making. National Research Council. National Academies Press, 2005.

Millennium Ecosystem Assessment: www.millenniumassessment.org/en/index.aspx

© 2007 SCIENTIFIC AMERICAN, INC.

When people talk about a moment being burned into memory, they usually mean it in a negative way: President John F. Kennedy's assassination, Princess Diana's fatal car crash, 9/11. The launch of Sputnik 50 years ago this month was different. It certainly had its negative side: no one likes to wake up to find that your nuclear adversary has thrown a shiny ball over your head and that you can't do a thing about it. But the dawn of the Space Age was also a hopeful event. Visionaries celebrated humanity's long-awaited climb out of its cradle, and pragmatists soon savored the benefits of communications and weather satellites. Many of today's scientists and engineers trace their life's passions to that fast-moving dot in the night sky. "In his millennia of looking at the stars, man has never faced so exciting a challenge as the year 1957 has suddenly thrust upon him," astronomers Fred L. Whipple and J. Allen Hynek wrote in the December 1957 issue of *Scientific American*.

THE FUTURE OF SPACE EXPLORATION

The launch of the Soviet Sputnik satellite half a century ago inaugurated the Space Age. What comes next?

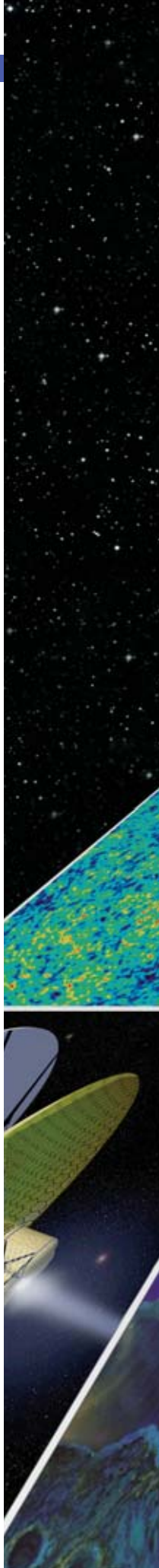
The evolution of the space program continues to be dramatic. In a decade or so, it will be hardly recognizable. The shuttle, which for all its faults is the most sophisticated flying machine ever built, will be a thing of the past. NASA is moving to the Constellation system, which is basically a high-tech dusting off of the Apollo rockets and capsules [see "To the Moon and Beyond," on page 62]. Whereas the shuttle is an ambitious spacecraft with modest goals (providing regular delivery-van service to orbit), Constellation is a modest spacecraft with ambitious goals: building a moon base, visiting an asteroid and eventually establishing human settlements on Mars. NASA Administrator Michael Griffin is steering a slow but steady course that he argues can be sustained on a limited budget—an approach that many commentators wish his predecessors had pursued 30 years ago.

For robotic missions to the planets, too, the coming decades promise to be transformative. Having completed the heroic initial reconnaissance of the solar system, the world's space agencies are digging deeper, in some cases literally. Perhaps the single biggest question, whether we are alone in the solar system, remains unanswered. Worlds once thought uninhabitable, such as the moons of Jupiter and Saturn, may have innards that are friendly to living things. By the centennial of Sputnik's launch, today's divide between human exploration and robotic science programs may have largely evaporated, as astronauts take on fieldwork too difficult for robots to conduct.

Both wings of the space program, human and robotic, badly need stability. The budgetary ups and downs of recent years, as well as swings in overall priorities, have resulted in dead ends and waste. In the second article of this special report, beginning on page 69, we set down scientists' top priorities in planetary science. Equally great missions in astronomy, cosmology and physics will be the subjects of *Scientific American* articles during the coming years.

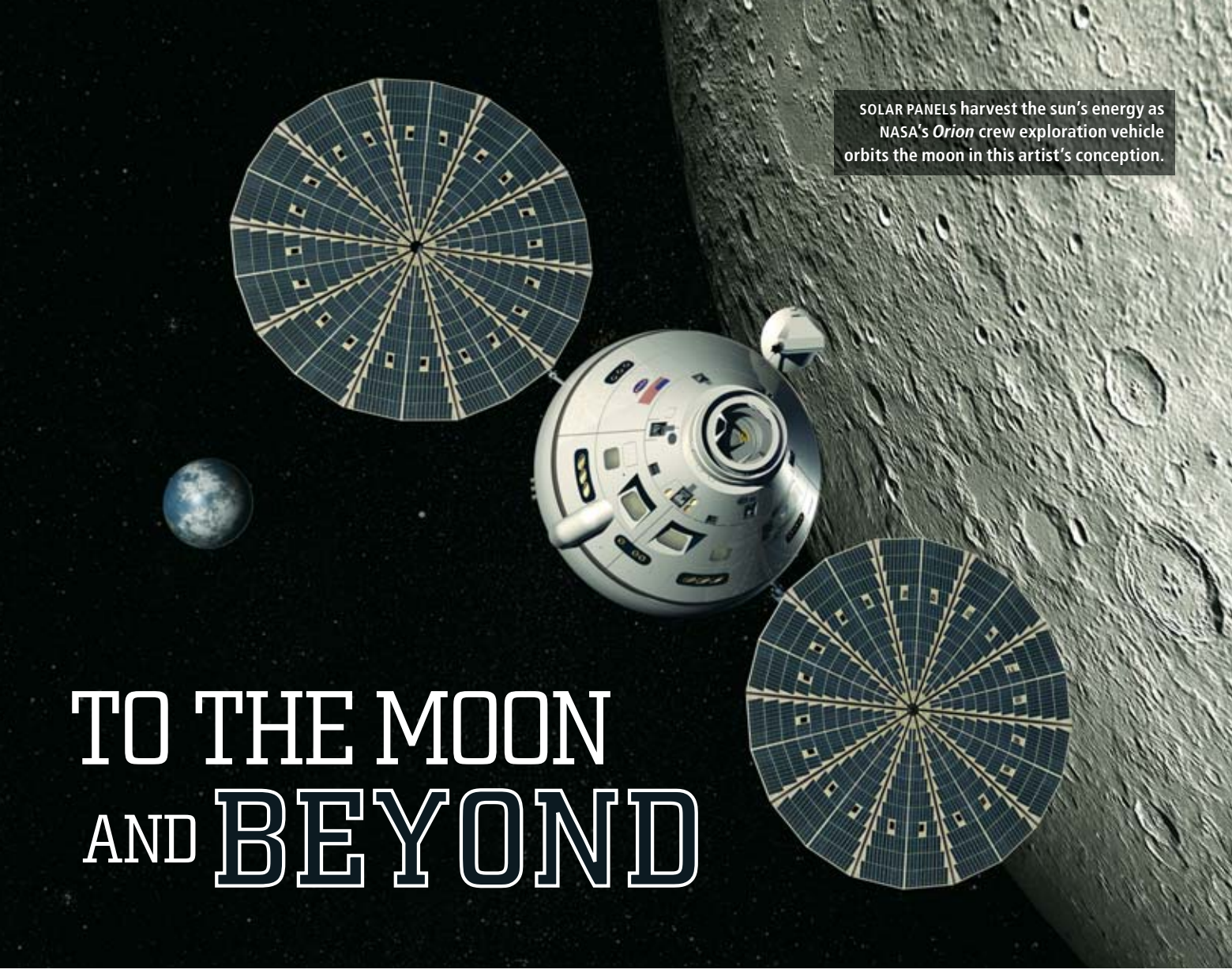
The world would be a different place but for the Space Age. Not all has been positive, but on balance, our lives are richer for it. By planning well now, we can ensure we will be able to say the same thing in 2057.

—Steven Ashley and George Musser, staff editors



© 2007 SCIENTIFIC AMERICAN, INC.

© 2007 SCIENTIFIC AMERICAN, INC.



SOLAR PANELS harvest the sun's energy as NASA's *Orion* crew exploration vehicle orbits the moon in this artist's conception.

TO THE MOON AND BEYOND

Humans are returning to the moon. This time the plan is to stay a while

By Charles Dingell, William A. Johns and Julie Kramer White

The moon, a luminous disk in the inky sky, appears suddenly above the broad crescent of Earth's horizon. The four astronauts in the *Orion* crew exploration vehicle have witnessed several such spectacular moonrises since their spacecraft reached orbit some 300 kilometers above the vast expanse of our home planet. But now, with a well-timed rocket boost, the pilot is ready to accelerate their vessel toward the distant target ahead. "Translunar injection burn in 10 seconds … " comes the call over the headset. "Five, four, three, two, one, mark … ignition…." White-hot flames erupt from a rocket nozzle far astern, and the entire ship—a stack of functional modules—vibrates as the crew starts the voyage to our nearest celestial neighbor, a still mysterious place that humans have not visited in nearly half a century. The year is 2020, and Americans are returning to the moon. This time, however, the goal is not just to come and go but to establish an outpost for a new generation of space explorers.

The *Orion* vehicle is a key component of the Constellation program, NASA's ambitious, multibillion-dollar effort to build a space transportation system that can not only bring humans to the moon and back but also resupply the International Space Station (ISS) and eventually place people on the planet Mars. Since the program was established in mid-2006, engineers and researchers at NASA, as well as at Lockheed Martin, *Orion*'s prime contractor, have been working to develop

STEPHEN C. HARTMAN

© 2007 SCIENTIFIC AMERICAN, INC.

the rocket launchers, crew and service modules, upper stages and landing systems necessary for the U.S. to mount a robust and affordable human spaceflight effort after its current launch workhorse, the space shuttle, retires in 2010.

To minimize development risks and costs, NASA planners based the Constellation program on many of the tried-and-true technical principles and know-how established during the Apollo program, an engineering feat that put men safely on the moon in the late 1960s and early 1970s. At the same time, NASA engineers are redesigning many systems and components using updated technology.

Orion starts with much the same general functionality as the *Apollo* spacecraft, and its crew capsule has a similar shape, but the resemblance is only skin-deep. *Orion* will, for example, accommodate larger crews than *Apollo* did. Four people will ride in a pressurized cabin with a volume of approximately 20 cubic meters for lunar missions (six will ride for visits to the space station starting around 2015), compared with *Apollo*'s three astronauts (plus equipment) in a cramped volume of about 10 cubic meters.

The latest structural designs, electronics, and computing and communications technologies will help project designers expand the new spacecraft's operational flexibility beyond that of *Apollo*. *Orion*, for instance, will be able to dock with other craft automatically and to loiter in lunar orbit for six months with no one onboard. Engineers are widening safety margins as well. In the event of an emergency during launch, for example, a powerful escape rocket will quickly remove the crew from danger, a benefit space shuttle astronauts do not enjoy. But to give you a better feel for what the program involves, let us start on the ground, before the *Orion* crew leaves Earth. From there, we will trace the progress of a prototypical lunar mission and the technologies planned to accomplish each stage.

Up, Up and Away

Towering 110 meters above the salt marshes of Florida's Kennedy Space Center, the two-stage Ares V cargo launch vehicle stands poised to blast off. The uncrewed vehicle, which contains a cluster of five powerful rocket engines, has almost the height and girth of the massive Saturn V rocket of *Apollo* fame. Derived from the space shuttle's external tank, Ares V's central booster tank delivers liquid-oxygen-hydrogen propellants to the vehicle's RS-68 engines—each a modified version of the ones currently used in the Delta IV military and commercial launcher. Two "strap-on," solid-fuel rocket boosters adapted from the space shuttle's system flank Ares V's central cylinder. They add the extra thrust that the launcher will need to loft the bug-like lunar lander and the "Earth departure stage"—a propulsion module that contains a liquid-oxygen-hydrogen-fueled J-2X engine (a descendant of NASA's Apollo-era Saturn V J-2 motor, built by Pratt & Whitney Rocketdyne) that will enable *Orion* to escape Earth's gravity and travel to the moon.

Abruptly, a flash exits the tail of the Ares V, and mounds of billowing smoke clouds soon envelop the booster, gantry and launchpad. After a momentary pause, a tremendous roar echoes across the spaceport, sending birds fleeing in all directions. Slowly at first, the big rocket ascends atop an ever expanding column of gray-white exhaust. Accelerating steadily, the vehicle blazes a smoky trail across the sky and disappears into the heavens. Minutes later, amid the silence of near-Earth space, Ares V jettisons its strap-on boosters, which fall into the sea, where they will be recovered. It then sheds the protective cargo sheath that covers its nose, revealing the lunar landing module. Circling the globe at an altitude of about 300 kilometers, the robot spacecraft now awaits the next step in the lunar excursion plan: rendezvous with *Orion*.

That same day the four moon-bound astronauts perch 98 meters above another Kennedy launchpad, anticipating imminent liftoff. Just below their conical *Orion* crew capsule is a drum-shaped service module that contains the spacecraft's on-orbit propulsion engine and much of its life-support system. Protective fairings envelop both to shield them from the strong aerodynamic forces and harsh conditions they

ORION BASICS

NASA and Lockheed Martin are developing a space transport system that, by the year 2020, will convey humans to and from the moon. The *Orion* vehicle—which contains a pressurized capsule, life-support systems and a propulsion engine—serves as one of the linchpins of the Constellation program, which also includes launch boosters and support modules.

Although the Constellation-Orion program in some ways resembles the 1960s-era Apollo moon program, it will be capable of conducting other manned and unmanned missions besides lunar excursions, including servicing the space station and perhaps traveling to the planet Mars.

The Constellation system's enhanced flexibility derives from a larger, more generalized design and the use of state-of-the-art technology.

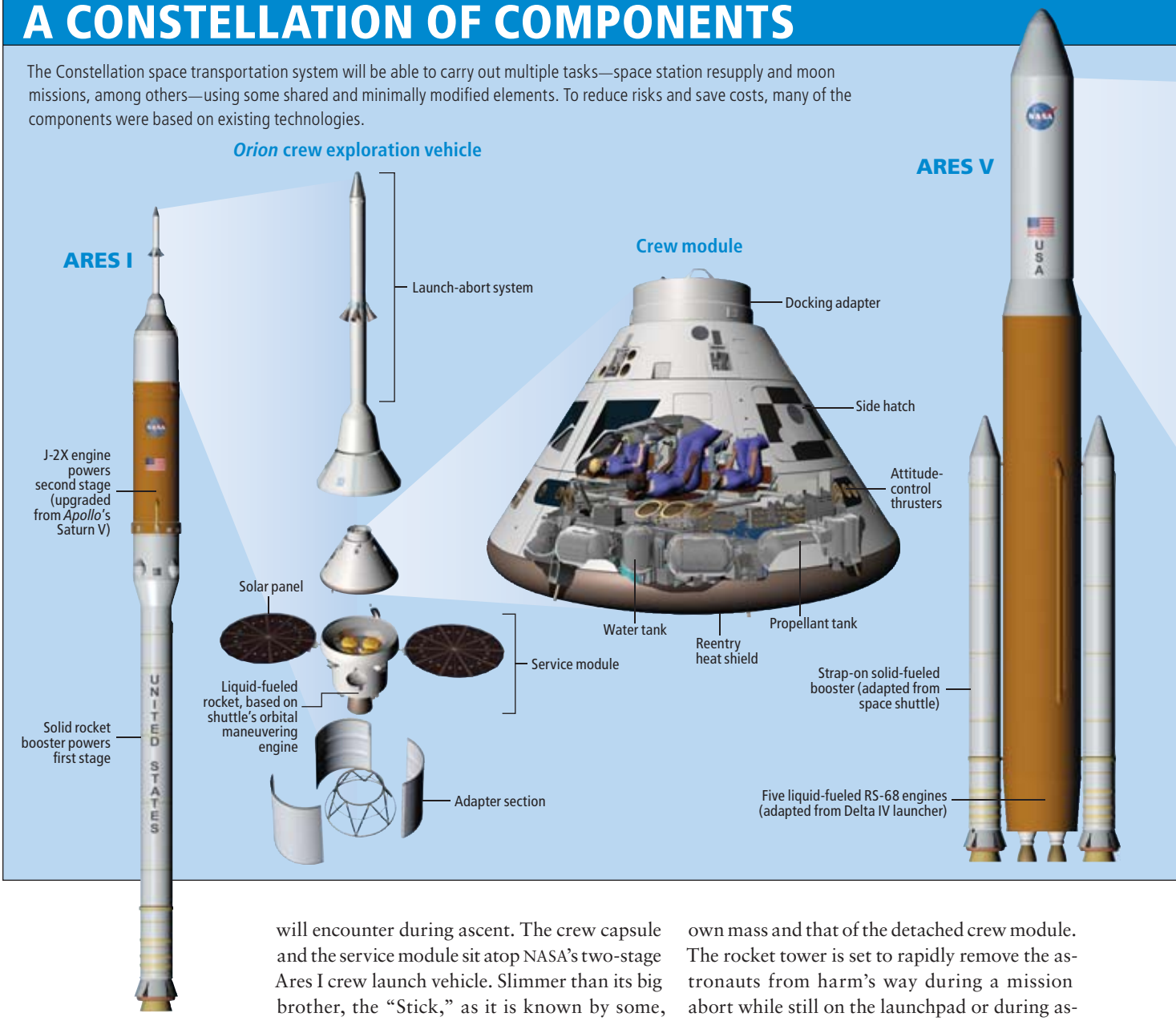
PAGE 61: ALL IMAGES COURTESY OF NASA EXCEPT AS NOTED: FAR RIGHT, SECOND ROW FROM TOP: U.S. NAVY (*Vanguard explosion*); SECOND LEFT, SIXTH ROW FROM TOP: ESA (*ExoMars*)

© 2007 SCIENTIFIC AMERICAN, INC.

[NASA'S NEW CREWED SPACECRAFT]

A CONSTELLATION OF COMPONENTS

The Constellation space transportation system will be able to carry out multiple tasks—space station resupply and moon missions, among others—using some shared and minimally modified elements. To reduce risks and save costs, many of the components were based on existing technologies.



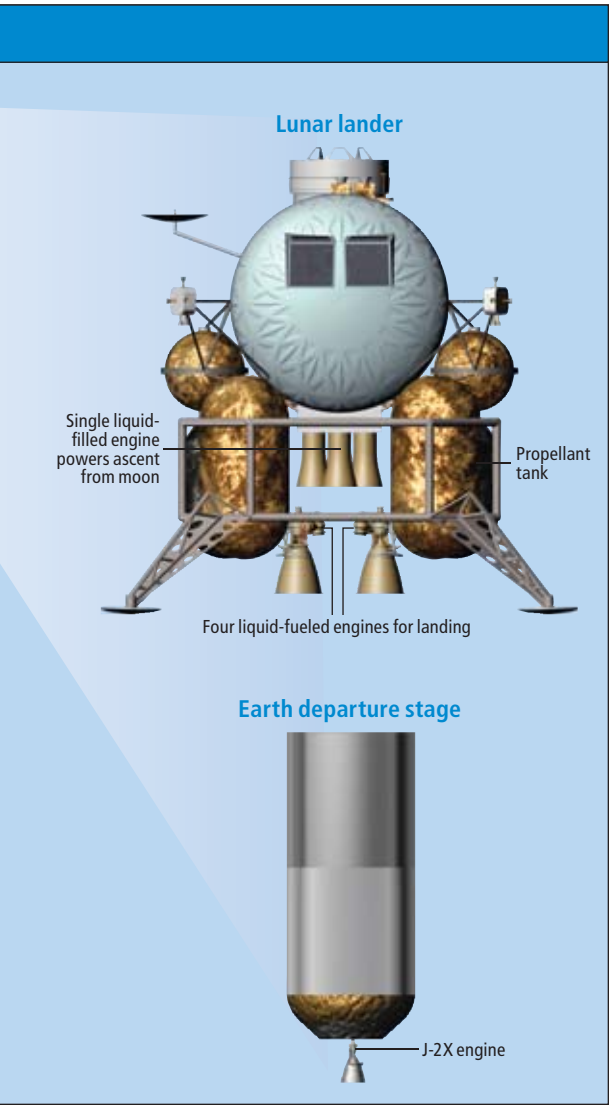
will encounter during ascent. The crew capsule and the service module sit atop NASA's two-stage Ares I crew launch vehicle. Slimmer than its big brother, the "Stick," as it is known by some, comprises another modified solid shuttle booster (constructed by Alliant Techsystems) topped by a second stage that is powered by a single J-2X motor. A spacecraft adapter serves as the structural and electrical interface between the *Orion* spacecraft and Ares I.

Capping the tall stack is an escape tower that is primed to rocket the occupants away from danger in the event of a failure. As the 1986 *Challenger* accident proved, space shuttle crews have little chance of survival if their ship sustains a major technical problem during launch and early ascent. *Orion*'s launch-abort system (LAS), in contrast, can for a few seconds impart a thrust that is equivalent to about 15 times its own mass and that of the detached crew module. The rocket tower is set to rapidly remove the astronauts from harm's way during a mission abort while still on the launchpad or during ascent. Should a serious glitch occur on the ground, the separated system would reach an altitude of about 1,200 meters to allow for parachute deployment and a downrange, or horizontal, distance of about 1,000 meters to clear the launchpad. Mission planners estimate that the LAS, together with *Orion*'s advanced guidance and control system, would be able to return the crew safely 999 out of 1,000 times it is needed.

But any such thoughts recede rapidly as the exhilaration of the impending launch mounts. As the countdown nears zero, commander and pilot intently eye the flight instruments on the flat-screen displays of *Orion*'s "glass cockpit," adapted from a safety-redundant version of the

STEPHEN C. HARTMAN

© 2007 SCIENTIFIC AMERICAN, INC.



avionics system used by advanced airliners such as the newly introduced Boeing 787 Dreamliner. The cockpit, with its computerized, fully electric "fly by wire" controls, energy-conserving electrical equipment and few mechanical switches, would be nearly unrecognizable to an Apollo-era astronaut.

A shudder ripples up through the entire structure, followed by a thunderous rumble. The Stick starts to move skyward. Gaining speed with every second, it rises rapidly, pressing the astronauts into their seats.

Almost two and a half minutes into the flight, the solid rocket booster is driving Ares I upward at a speed of Mach 6. At a height of about 61,000 meters, the first stage separates and falls back to Earth on parachutes so that it may be recovered and later recycled. Meanwhile the J-2X second-stage rocket motor ignites, sending the *Orion* crew module, the service module and the LAS through the last reaches of the atmosphere. Their usefulness ended now that the craft has exited the atmosphere, the aerodynamic shrouds break away to maximize ascent performance by shedding weight. By this time the vessel has gained enough velocity to reduce the risk of an emergency abort, so the LAS and its protective fairing also separate and fall away. The second-stage engine cuts off as the crew capsule and the service module near an altitude of about 100 kilometers.

Rendezvous in Earth Orbit

The service module engine then ignites, completing the job of inserting *Orion* into orbit and initiating the maneuvers it needs to rendezvous with the Earth departure stage and the lunar lander. *Orion*'s main engine is adapted from the flight-proved space shuttle orbital maneuvering engine, upgraded for greater propulsion thrust and efficiency. The service module contains power generation and storage systems, radiators that expel surplus heat into space, all necessary fluids and a science equipment bay. To maximize space in the crew vehicle, the service module also carries some of the avionics system, as well as part of the environmental control and life-support subsystems. A lightweight polymer-composite honeycomb reinforced with aluminum forms its structure; simple manufacturing methods should help keep down the cost of this expendable item.

One of the more notable differences between *Orion* and *Apollo* is the addition to the service module of umbrella-shaped solar arrays that unfold when needed in orbit. Because the *Apollo* spacecraft was designed for moon missions measured in days, it carried hydrogen fuel cells that could generate electrical power only for relatively short periods. *Orion,* in contrast, must be able to produce electricity for at least six months.

Gradually, *Orion* catches up to the lunar lander and departure stage that Ares V had earlier placed into low Earth orbit. When the two craft finally rendezvous, the crew performs (or monitors) the final maneuvers and keeps an eye on the automated "soft capture" system as it aligns the pair and then smoothly docks them. Force-feedback and electromechanical components sense loads, automatically capturing the mating rings of the vehicles and actively damping out any contact forces. Ship and crew are now nearly ready to head for the moon.

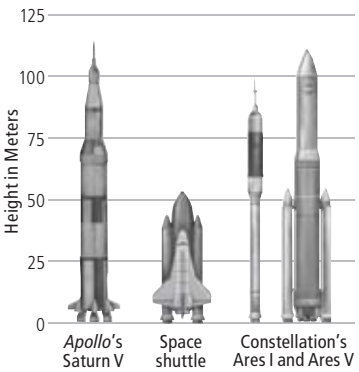
The crew module is the only element of *Orion* that will make the entire trip, and it may be

[THE AUTHORS]

Charles Dingell, William A. Johns and Julie Kramer White manage engineering and technology operations for the NASA/Lockheed Martin Orion project. Dingell serves as the project's technical director for NASA. During a quarter century at the agency, he held leadership positions supporting the space shuttle, X-38 crew return vehicle and orbital space plane programs. Johns is chief engineer and technical director of the *Orion* crew exploration vehicle (CEV) for Lockheed Martin, the lead contractor on the project. After starting at Martin Marietta in 1980, he worked on successive versions of the Centaur upper stage and later led development of the Atlas V evolved expendable launch vehicle. Kramer White is NASA's chief engineer for the *Orion* CEV project. She has more than 20 years of technical management experience on the space shuttle, space station and X-38 programs.

LAUNCH SYSTEM LINEUP

NASA has fielded three orbital launch systems for crewed missions since the late 1960s.



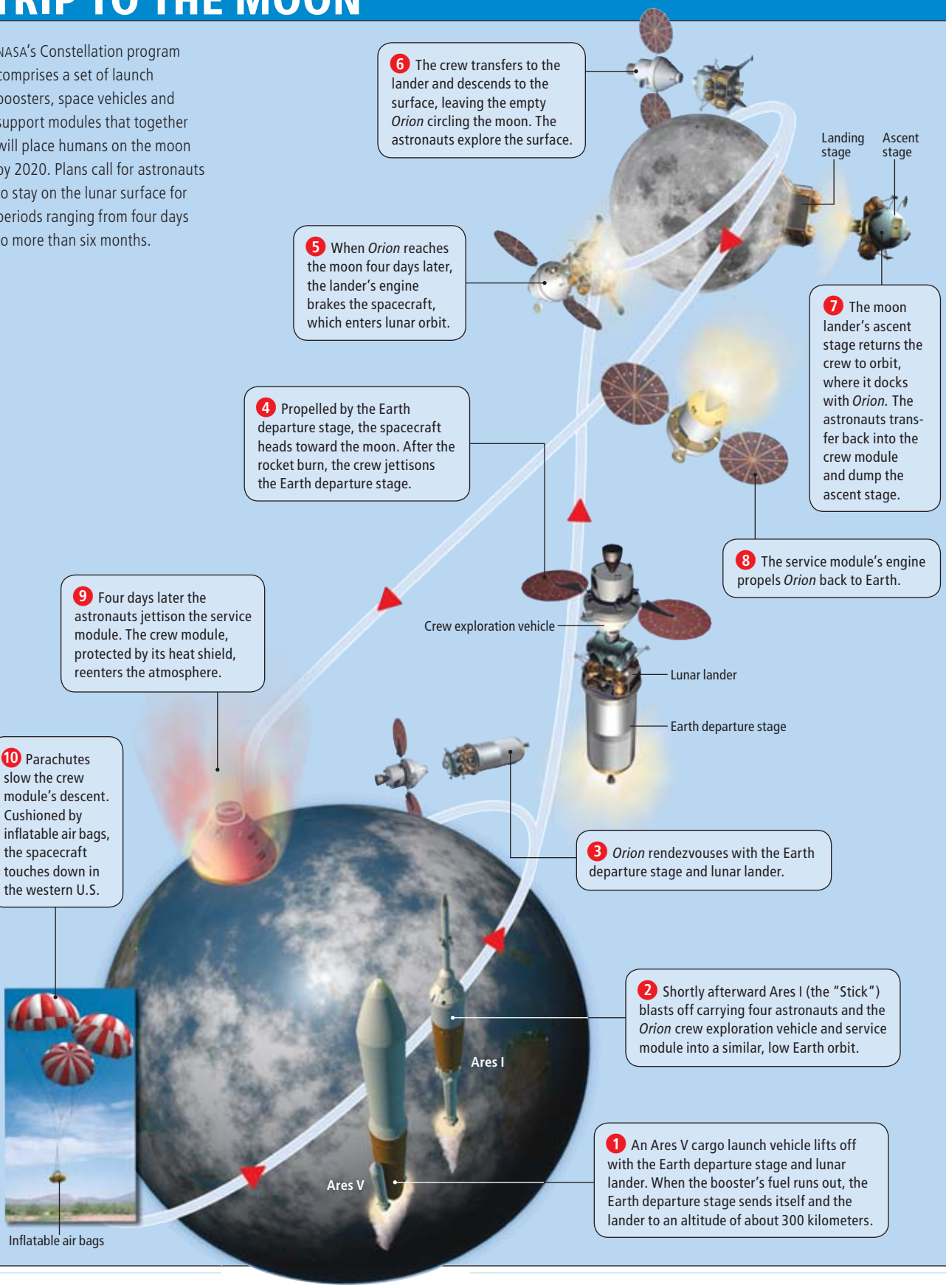
ANN SANDERSON (*graph*)

© 2007 SCIENTIFIC AMERICAN, INC.

[SAMPLE MISSION: STEP BY STEP]

TRIP TO THE MOON

NASA's Constellation program comprises a set of launch boosters, space vehicles and support modules that together will place humans on the moon by 2020. Plans call for astronauts to stay on the lunar surface for periods ranging from four days to more than six months.



STEPHEN C. HARTMAN

© 2007 SCIENTIFIC AMERICAN, INC.

reused for up to 10 flights. A lightweight aluminum-lithium alloy with titanium reinforcements makes up most of the capsule structure. The exterior of the crew vehicle is lined with a thermal protection system, which, in addition to protecting its living quarters from the searing heat of reentry, also incorporates a tough, impact-resistant layer that shields it against high-velocity micrometeoroids or other debris that may strike its outer surface.

The crew module's reaction-control maneuvering system uses gaseous oxygen and methane propellants, a technology that builds on the progress engineers made during NASA's X-33 single-stage-to-orbit vehicle program, which was canceled in 2001. One advantage of the oxygen-methane propulsion system is that its fuel will be nontoxic (unlike its predecessors that used hypergolic propellants), which will help ensure the safety of the flight and ground crews after they return to Earth.

When all is ready, the Earth departure stage rocket engine ignites to propel the spacecraft toward the moon. Engineers are configuring *Orion* to support both "lunar sortie missions," in which crew members spend four to seven days on the moon's surface to demonstrate the *Orion* system's ability to transport and land humans on Earth's satellite, and "lunar outpost missions," in which a semicontinuous human presence would be established there. Because the maximum duration of a crew's stay on the lunar surface is 210 days (determined by the available supplies of oxygen, water and other consumables), *Orion*'s continuous operation capability must exceed that period. The biggest design driver for *Orion* lunar missions is the amount of propellant required to meet these objectives.

After a four-day trip outbound, the crew enters into lunar orbit, having dumped the Earth departure stage along the way. The four astronauts climb into the lander, leaving the crew capsule and service module to wait for them in orbit. As with the *Apollo* lunar excursion module, the lunar lander consists of two components. One is the descent stage, which has legs to support the craft on the surface as well as most of the crew's consumables and scientific equipment. The other part is the ascent stage that houses the crew. After landing and exploring the surface, the foursome blasts off the moon's surface and later docks with the crew and service modules in orbit. The ascent stage of the lander is discarded into outer space, and *Orion* rockets back to Earth.

Return to the Home Planet

As the *Orion* astronauts close in on the blue planet, they may have to prepare for a reentry and landing quite unlike those of *Apollo*. Like the *Gemini* and *Mercury* spacecraft before it, *Apollo* splashed down in the ocean after it had plunged through the atmosphere. But because water landings would require costly fleets of recovery ships and expose a reusable spacecraft to saltwater corrosion, NASA planners may decide that *Orion* should touch down on land, as the Russian *Soyuz* spacecraft does. *Orion*'s greater size, weight and lift, however, exacerbate the engineering challenge. The "land landing" mode is also important to minimizing life-cycle costs. If the agency instead opts to land in the ocean, *Orion* will be fitted with much the same capabilities as *Apollo*.

Unfortunately, setting down on American soil after a lunar mission presents a fundamental problem. For nearly half of the lunar month, orbital conditions would place any landing site in the Southern Hemisphere, away from the planned locations in the western continental U.S. Although the time of departure from lunar orbit can vary the longitude of the reentry point, its latitude is fixed by the declination (angular distance from the equator) of the moon relative to Earth at lunar departure. Thus, to reach landing sites in the western U.S. or waters near the continental U.S. during unfavorable periods of the lunar month, *Orion* will stretch its land-

ASTRONAUT ASTROPHYSICS

NASA has selected four concept studies that may lead to experiments that will be set up by* Orion *astronauts when they reach the lunar surface:

Two concept studies proposed separately by researchers at the University of Maryland at College Park and the NASA Goddard Space Flight Center are designed to determine the distance from Earth to the moon with submillimeter accuracy.

A scientist at the Naval Research Laboratory plans a radio-telescope array on the far side of the moon to study elementary particle acceleration in supernovae, quasars and the solar corona.

"Soft" x-ray emissions produced by interactions between the solar wind and Earth's magnetic field and by our galaxy's core are to be measured by a package developed by a Goddard investigator.

ENGINEERING CHALLENGES

Air-bag drop test

- SOFT LANDING SYSTEM: Project engineers are developing an air-bag system (*above*) that inflates underneath the *Orion* crew module to cushion the shock of landing on hard terrain, if the vehicle should do so. If the air-bag technology proves unsuitable, however, they have several alternatives in the works, including retrorocket systems.
- OVERALL TRADE-OFFS: Engineers must find the optimal balance between the maximum payload masses that the Ares boosters can deliver, on the one hand, and the combination of the widest safety margins, greatest system redundancy and most useful crew accommodations on the other.
- HEAT SHIELD: Although the engineers favor PICA (phenolic impregnated carbon ablator) as the capsule's heat-shield material, building such a large structure could pose problems. Heat-resistant space shuttle tiles or *Apollo*'s ablator material could serve as substitutes.

NASA LANGLEY RESEARCH CENTER

© 2007 SCIENTIFIC AMERICAN, INC.

ing point into the Northern Hemisphere by employing aerodynamic lift produced as it descends into Earth's outer atmosphere. A trajectory of this type, in which a spacecraft bounces across the upper atmosphere like a stone skipping across a pond, is sometimes known as a skip reentry.

Having spent the four-day return journey from the moon fine-tuning *Orion*'s flight path for the first crewed skip-reentry maneuver ever attempted, anticipation builds among the astronauts as the blue-white visage of our home planet grows ever larger in their view screen. They are soon occupied, however, by reorienting the ship so that the service module can be jettisoned, a necessary operation that exposes the protective heat shield on the crew module's underside. Later, after using *Orion*'s redundant navigation system and flight computers to check that the spacecraft's attitude is positioned properly for reentry and that its trajectory is following the correct, shallow-angle route, the crew prepares for the onset of deceleration forces as *Orion* encounters the atmosphere.

The skip-reentry process starts out slowly. At first, the crew begins to notice weak g-forces caused by the resistance of the thin, high-altitude air. The g-forces, which push the crew members against their seats, grow steadily in strength as bits of glowing heat-shield material and streams of ionized gas streak past the windows. Shortly after *Orion* starts to scrape against the upper reaches of the atmosphere, the spacecraft rebounds briefly to a higher altitude.

▼ELECTRIC ARC—essentially a room-size blowtorch—blasts air heated to several thousand degrees onto a sample of experimental reentry heat-shield material at the NASA Ames Research Center.

After the skip, the capsule dives deeply into the air on a path toward the landing site.

The tragic loss of the *Columbia* space shuttle and crew in 2003 demonstrated that the thermal protection system of a returning vehicle is critical. Atmospheric reentry generates tremendous heating on the undersurface of spacecraft (a couple of thousand degrees Celsius) caused by the friction of the air rushing by at hypersonic speeds. Because *Orion*'s reentry velocity from a moon mission (which is on the order of 11 kilometers a second) will be 41 percent faster than a shuttle's descent speed from low Earth orbit, the heat load will be several times greater. The fact that the *Orion* crew module is larger than that of *Apollo* compounds the challenge.

The leading candidate for *Orion*'s base heat shield is a material called PICA (phenolic impregnated carbon ablator). PICA is a matrix of carbon fibers embedded in a phenolic resin. At high temperatures, the outer surface of the PICA layer ablates, or burns away, to carry off much of that extreme heat. The ablator's surface pyrolyzes when heated, leaving a heat-resistant layer of charred material. PICA's low thermal conductivity also blocks heat transfer to the crew module. PICA was used in 2006, when it protected the Stardust spacecraft (which carried a sample from Comet Wild 2) as it came back to Earth at 13,000 meters a second—the fastest controlled reentry ever. Being 40 times larger in area, *Orion*'s heat shield will need to be built in segments, thus adding new complexities.

Landing on Land

Finally, three large parachutes—which closely resemble those used by *Apollo*—deploy to slow the vehicle's rate of descent. The reassuring sight of the voluminous red-and-white canopies opening above tells the astronauts that their amazing trip is almost complete. Before long, *Orion* is jarred by the release of its large heat shield. Hanging below the big chutes, the crew module now descends at about eight meters a second.

In the case of a "land landing," an airbag system inflates on the crew module's underside to absorb and attenuate the upcoming landing shock. With a solid jolt, the spacecraft at last sets down on dry land in the western American desert. *Orion* has returned home. ■

SCI AM

NASA and Lockheed Martin have prepared a series of photographs and conceptual illustrations related to the Constellation/Orion missions. Go to **www.SciAm.com/ontheweb**

MORE TO EXPLORE

Lockheed Martin *Orion* crew vehicle Web page: **www.lockheedmartin.com/wms/findPage.do?dsp=fec&ci=17675**

NASA Constellation program Web site: **www.nasa.gov/mission_pages/constellation/main/index.html**

NASA Vision for Space Exploration Web site: **www.nasa.gov/mission_pages/exploration/main/index.html**

NASA AMES RESEARCH CENTER

© 2007 SCIENTIFIC AMERICAN, INC.

5 ESSENTIAL THINGS TO DO IN SPACE

Planetary scientists have articulated goals for exploring the solar system

By George Musser

TO A CHILD OF THE SPACE AGE, books about the solar system from before 1957 are vaguely horrifying. How little people knew. They had no idea of the great volcanoes and canyons of Mars, which make Mount Everest look like a worn hillock and the Grand Canyon like a roadside ditch. They speculated that Venus beneath its clouds was a lush, misty jungle, or maybe a dry, barren desert, or a seltzer water ocean, or a giant tar pit—almost everything, it seems, but what it really is: an epic volcanic wasteland, the scene of a Noah's flood in molten rock. Pictures of Saturn were just sad: two fuzzy rings where today we see hundreds of thousands of fine ringlets. The giant planet's moons were gnats, rather than gnarled landscapes of methane lakes and dusty geysers.

All in all, the planets seemed like pretty small places back then, little more than smudges of light. At the same time, Earth seemed a lot larger than it does now. No one had ever seen our planet as a planet: a blue marble on black velvet, coated with a fragile veneer of water and air. No one knew that the moon was born in an impact or that the dinosaurs died in one. No one fully appreciated that humanity was becoming a geologic force in its own right, capable of changing the environment on a global scale. Whatever else the Space Age has done, it has enriched our view of the natural world and given us a perspective that we now take for granted.

Since Sputnik, planetary exploration has gone through several waxing and waning phases. The 1980s, for instance, might as well have been the dark side of the moon. The present looks brighter: dozens of probes from the world's space programs have fanned out across the solar system, from Mercury to Pluto. But budget cuts, cost overruns and inconsistency of purpose have cast long shadows over NASA. At the very least the agency is going through its most unsettled period of transition since Nixon shot down the Apollo moon missions 35 years ago.

"NASA continues to wrestle with its own identity," says Anthony Janetos of the Pacific Northwest National Laboratory, a member of a National Research Council (NRC) panel that scrutinized NASA's Earth observation program. "Is it about exploring space? Is it about human exploration, is it about science, is it about exploring the outer universe, is it about exploring the solar system, is it about the space shuttle and station, is it about understanding this planet?"

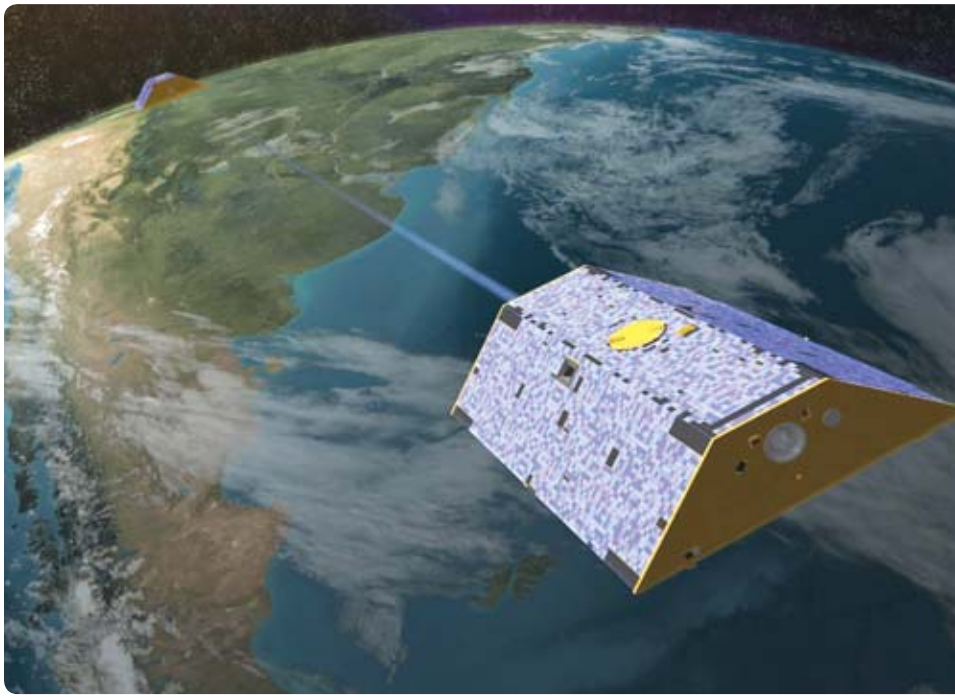
In principle, the upheaval should be a happy occasion. Not only are robotic probes flying hither and yon, the human space program is no longer drifting like a spent rocket booster. President George W. Bush set out a clear and compelling goal in 2004—namely, to plant boots in lunar and Martian soil. Though controversial, the vision gave NASA something to shoot for. The trouble is that it quickly turned into an unfunded mandate, forcing the agency to breach the "firewall" that had traditionally (if imperfectly) shielded the science and human spaceflight programs from each other's cost overruns. "I presume it is not news to you that NASA doesn't have enough money to do all the things it's being asked to do," says Bill Claybaugh, director of NASA's Studies and Analysis Division. Cash doesn't exactly flow like liquid hydrogen at space agencies in other countries, either.

NRC panels periodically take a step back and ask whether the world's planetary exploration programs are on track. The list of goals that follows synthesizes their priorities.

NASA, ESA, AND J. HESTER AND A. LOLL *Arizona State University*

© 2007 SCIENTIFIC AMERICAN, INC.

1 Monitor Earth's Climate

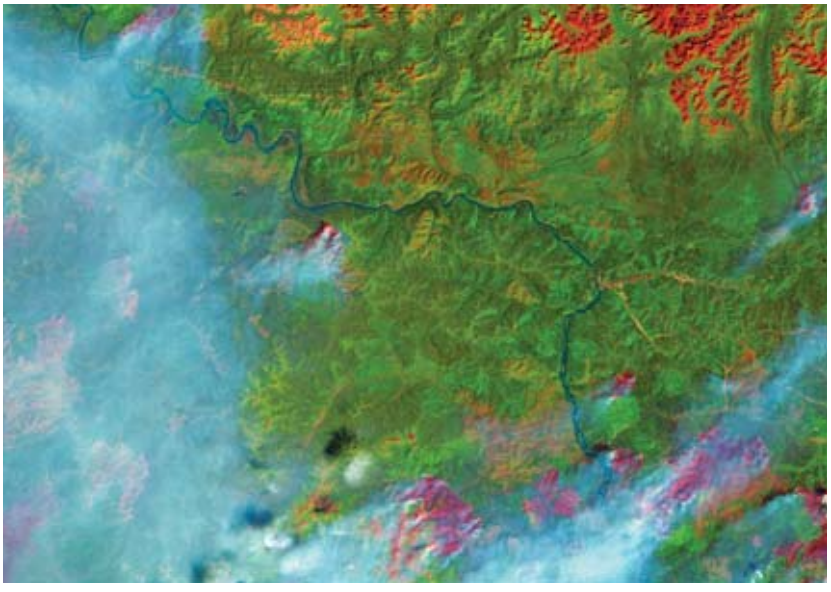


▲ TWIN SATELLITES of the Gravity Recovery and Climate Experiment (GRACE), which detect the gravitational distortions caused by the movement of water, are already past their original planned lifetime.

Amid all the excitement of buggying around Mars and peeling back the veil of Titan, people sometimes take the mundane yet urgent task of looking after our own planet for granted. NASA and the National Oceanic and Atmospheric Administration (NOAA) have really let it slide. In 2005 Janetos's NRC panel argued that the "system of environmental satellites is at risk of collapse." The situation then deteriorated further. NASA shifted $600 million over five years from Earth science to the shuttle and space station. Meanwhile the construction of the next-generation National Polar-Orbiting Operational Environmental Satellite System ran seriously over budget and had to be downsized, stripping out instruments crucial to assessing global warming, such as those that measure incoming solar radiation and outgoing infrared radiation.

Consequently, the two dozen satellites of the Earth Observation System are reaching the end of their expected lifetimes before their replacements are ready. Scientists and engineers think they can keep the satellites going, but there is a limit. "We could hold out, but we need a plan now," says Robert Cahalan, head of the Climate and Radiation Branch at the NASA Goddard Space Flight Center. "You can't wait till it breaks."

If a satellite dies before relief arrives, gaps open up in the data record, making it difficult to establish trends. For instance, if a newer instrument discovers that the sun is brighter than its predecessor found, is it because the sun really brightened or because one of the instruments was improperly calibrated? Unless satellites overlap in time, scientists may not be able to tell the difference. The venerable Landsat series, which has monitored the surface since 1972, has been on the fritz for years, and the U.S. Department of Agriculture has already had to buy data from Indian satellites to monitor crop productivity. For some types of data, no other nation can fill in.



▼ LANDSAT 7 IMAGE, taken at multiple wavelengths, shows fires in Alaska and the Yukon in 2004. A failure in 2003 hobbled the satellite, and the entire program has been on budgetary drip-feed for more than a decade.

The NRC panel called for restoring the lost funding, which would pay for 17 new missions over the coming decade, such as ones to keep tabs on ice sheets and carbon dioxide levels—essential for predicting climate change and its effects. The root issue, though, is that climate observations fall somewhere in between routine weather monitoring (NOAA's specialty) and cutting-edge science (NASA's). "There's a fundamental problem that no one is charged with climate monitoring," says climatologist Drew Shindell of the NASA Goddard Institute for Space Studies. He and others have suggested that the U.S. government's scattered climate programs be consolidated in a dedicated agency, which would own the problem and give it the focus it deserves.

ACTION PLAN

- Fund the 17 new satellites proposed by the National Research Council over the next decade (estimated cost: $500 million a year)
- Found a climate agency

NASA/JPL (*top*); IMAGE BY JESSE ALLEN, BASED ON DATA FROM USGS EROS DATA CENTER (*bottom*)

© 2007 SCIENTIFIC AMERICAN, INC.

2 Prepare an Asteroid Defense

▲ NOT A SIGHT YOU WANT TO SEE: An asteroid of the dinosaur-killing sort would reach from sea level to the cruising altitude of commercial airliners.

Like climate monitoring, guarding the planet from asteroids always seems to fall between the cracks. Neither NASA nor the European Space Agency (ESA) has a mandate to stave off human extinction. The closest they come is NASA's Spaceguard Survey, a $4-million-a-year telescope observing program to scan near-Earth space for kilometer-size bodies, the range that can cause global as opposed to merely regional destruction. But no one has done a systematic search for region destroyers, an estimated 20,000 of which come within striking range of our planet. No Office of Big Space Rocks is standing by to evaluate threats and pick up the red phone if need be. It would take 15 years or longer to mount a defense against an incoming body, assuming that the technology were ready to go, which it isn't. "Right now the U.S. doesn't have a comprehensive plan," says aerospace engineer Larry Lemke of the NASA Ames Research Center.

This past March, at Congress's request, NASA published a report that could serve as the starting point for such a plan. By its analysis, searching for 100- to 1,000-meter bodies could largely piggyback on the Large Synoptic Survey Telescope (LSST), an instrument that a consortium of astronomers and companies (most famously Google) is now working on to scan the sky for anything that moves, blinks or winks. A report last month by the LSST project itself estimates that the scope, as currently designed, should find 80 percent of the bodies over one decade of operation, from 2014 to 2024. With an extra $100 million of fine-tuning, it could net 90 percent of them.

Like any Earth-based instrument, though, the LSST has two limitations. First, it has a blind spot: bodies that are either just ahead or just behind Earth in its orbit—the most dangerous ones—can be observed only at dusk or dawn, when they are easily lost in the sun's glare. Second, the instrument can estimate the mass of asteroids only indirectly, based on how bright they are. Limited to visible light, the estimates are good only to a factor of two: a big but dark asteroid can masquerade as a small but bright one. "That difference could matter a lot if we actually decided there was a mitigation required," NASA's Claybaugh says.

THREATENING ASTEROIDS

Dinosaur killers, 10 kilometers across, hit every 100 million years on average. Globally devastating asteroids, one kilometer or larger, come every half a million years or so. City destroyers, 50 meters across, strike perhaps once a millennium.

The Spaceguard Survey has found just over 700 kilometer-size bodies, none posing a threat over the coming centuries. The rate of discovery is tapering off, suggesting the survey has found about 75 percent of the total.

The chance of an impactor among the remaining 25 percent is small, but the consequence would be large. Statistically, the risk amounts to an average death toll of 1,000 people a year. Smaller ones kill maybe 100 a year on average.

To plug these holes, the NASA team also considered building a $500-million infrared space telescope and putting it in its own orbit around the sun. It could pick up essentially every threat to Earth and, by studying bodies at multiple wavelengths, pin down their mass to within 20 percent. "If you want to do it right, you want to go to the infrared in space," says planetary scientist Donald Yeomans of the Jet Propulsion Laboratory (JPL), a co-author of the report.

The other question is what to do if an asteroid is on its way. A rule of thumb is that to divert an asteroid by one Earth radius, you need to change its velocity by one millimeter per second, one decade in advance, by either hitting it, nuking it, pushing it or towing it gravitationally. In 2004 ESA's Near-Earth Object Mission Advisory Panel recommended doing a trial run. Known as Don Quijote, the proposed $400-million mission would fire a 400-kilogram projectile into an asteroid and watch what happens.

The debris thrown out by the impact would exert a force on the asteroid via the rocket effect, but no one knows how strong it would be. Finding out is the whole point of the mission. "You can find out whether the kinetic impactor strategy would work or not," says Alan Harris of the German Aerospace Center in Berlin. Scientists would choose a body on a distant orbit so that

◀ THY ASTEROID IS SO GREAT, and my spaceship is so small. But given enough time, even a modest rocket can steer a big rock out of harm's way.

REUTERS/CORBIS (*top*); PAT RAWLINGS *SAIC* (*bottom*)

© 2007 SCIENTIFIC AMERICAN, INC.

a snafu could not end up putting it on a collision course with Earth.

This past spring ESA completed a set of feasibility studies—and promptly shelved them for lack of money. It would take a joint effort with NASA or the Japan Aerospace Exploration Agency (JAXA), or both, to make the plan happen.

ACTION PLAN

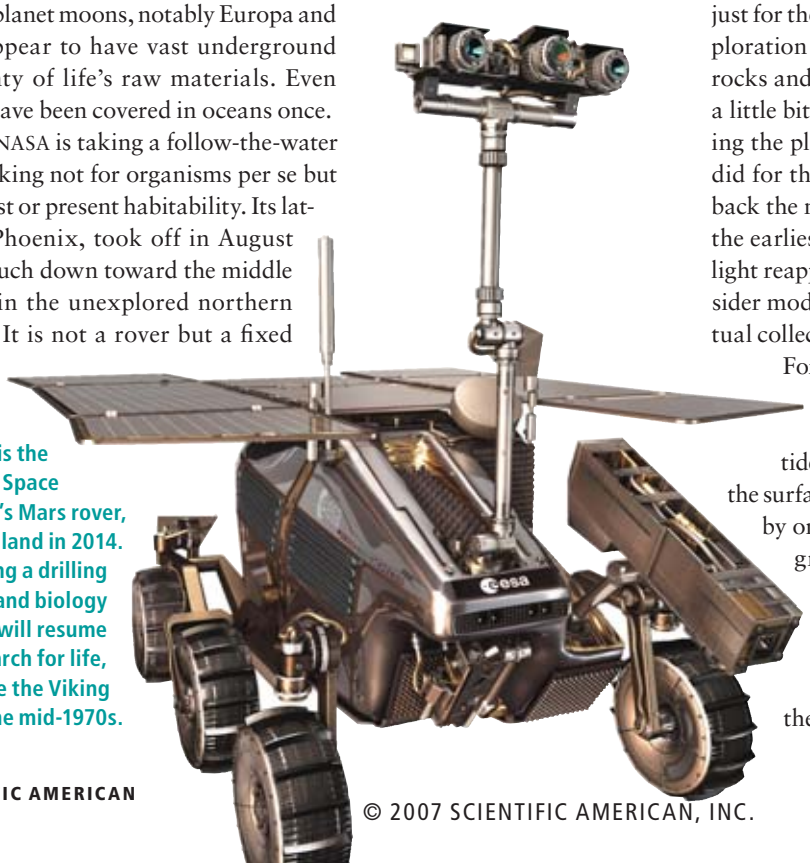
- Extend asteroid search to smaller bodies, perhaps using a dedicated infrared space telescope
- Deflect an asteroid in a controlled way as a trial run
- Develop an official system for evaluating potential threats

▲ HOT-AIR BALLOON is ideally suited to getting around Titan. A plutonium power source throws off enough waste heat for a 12-meter balloon carrying 160 kilograms of instruments and cruising at an altitude of 10 kilometers.

3 Seek Out New Life

Before Sputnik, scientists thought the solar system might be a veritable Garden of Eden. Then came the fall. Earth's sister worlds proved to be hellish. Even Mars bit the dust when the Mariner probes revealed a cratered moonscape and the Viking landers failed to find even a single organic molecule. But lately the plausible venues for life have multiplied. Mars is looking hopeful again. Outer-planet moons, notably Europa and Enceladus, appear to have vast underground seas and plenty of life's raw materials. Even Venus might have been covered in oceans once.

For Mars, NASA is taking a follow-the-water approach, looking not for organisms per se but for signs of past or present habitability. Its latest mission, Phoenix, took off in August and should touch down toward the middle of next year in the unexplored northern polar region. It is not a rover but a fixed lander with a robot arm capable of digging down a few centimeters to study shallow ice deposits. The agency's next step is the $1.5-billion Mars Science Laboratory (MSL), a car-size rover scheduled for launch in late 2009 and landfall a year later.



▶ EXOMARS is the European Space Agency's Mars rover, scheduled to land in 2014. Carrying a drilling platform and biology laboratory, it will resume the direct search for life, on hold since the Viking landers of the mid-1970s.

Gradually, though, scientists will want to return to the direct search for living things or their remains. In 2013 ESA plans to launch the ExoMars rover, equipped with a Viking-like chemistry lab and, crucially, a drilling platform able to go two meters down—which should be deep enough to get past the toxic surface layers to where organic material might have survived.

Unfortunately, the trail then goes cold. Most planetary scientists' single highest priority—not just for the search for life but for solar system exploration generally—is to bring some Martian rocks and dirt back to Earth for analysis. Even a little bit would go a long way toward unraveling the planet's history, as the Apollo samples did for the moon. NASA's budget woes pushed back the multibillion-dollar mission to 2024 at the earliest, but over the summer a glimmer of light reappeared when the agency began to consider modifying MSL to store samples for eventual collection.

For Europa, scientists' priority is an orbiter to measure how the satellite's shape and gravitational field respond to tides raised by Jupiter. If a sea lies within, the surface will rise and fall by 30 meters; if not, by only one meter. Magnetic readings and ground-penetrating radar could also dowse the ocean, and cameras would map the surface in preparation for an eventual lander and driller.

For Titan, a natural follow-up to the ongoing Cassini mission would be an

TIBOR BALINT (*top*); ESA (*bottom*)

© 2007 SCIENTIFIC AMERICAN, INC.

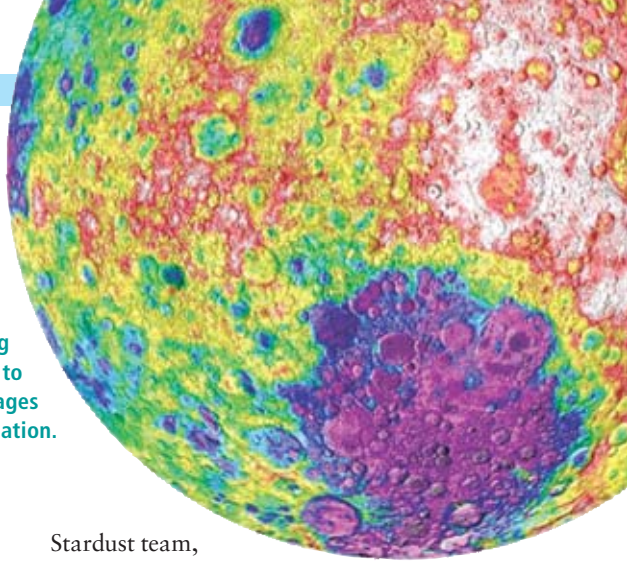
orbiter plus a surface sampler. Titan's Earth-like atmosphere opens up the possibility of a hot-air balloon, which could dip down every now and then to grab rocks and dirt. The goal, says Jonathan Lunine of the University of Arizona, would be to "analyze the surface organics to see if there are systematic trends that suggest the start of self-organization, which is how most origins-of-life people think life began on Earth."

This past January, NASA finally began to study these missions seriously. The agency plans to choose between Europa and Titan next year, so that a $2-billion probe could fly in about a decade. The body that doesn't make the cut will have to wait yet another decade.

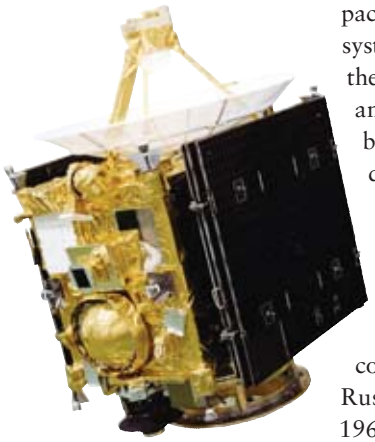
In the end, it may turn out that life on Earth is unique after all. Disappointing, no doubt, but it would not mean the whole effort had been a waste. "I see astrobiology more broadly than just looking for life," says Bruce Jakosky, director of the Center for Astrobiology at the University of Colorado. It is also about figuring out how varied life can or cannot be, what its preconditions are, and how lifelessness begat life four billion years ago on our planet. Thus, the search is not just about finding companionship in the cosmos. It is about divining our own origins.

ACTION PLAN

- Get Martian sample return on track
- Gear up for returning to Europa and Titan



▶ AITKEN BASIN (*purplish blotch*) on Earth's moon, considered the biggest hole in the solar system, is 2,500 kilometers across and 12 kilometers deep. Pinning down its age is crucial to unraveling the late stages of planet formation.



▲ HAYABUSA ASTEROID SAMPLER is an innovative (if trouble-plagued) probe that could serve as a model for a comet nucleus sample return mission—which planet formation experts consider crucial. ▼

4 Explain the Genesis of the Planets

Like the origin of life, the origin of the planets was a complex, multistage process. Jupiter was the first-born and the guiding hand for the rest. Did it build up slowly, like the other planets, or did it take shape in a single gravitational whoosh, like a small star? Did it form farther from the sun and move inward, as its anomalously high levels of heavy elements suggest—in which case it might have flicked lesser worlds out of its way? NASA's Juno orbiter to the giant planet, scheduled for launch in 2011, might provide some answers.

Those concerned with planet formation also want to follow up the Stardust mission, which returned samples last year from the coma of dust that surrounds a solid comet nucleus. "We have just scratched the surface," says the head of the Stardust team, Donald Brownlee of the University of Washington. "Stardust showed that comets were terrific collectors of early solar system materials from the entire solar nebula. These materials were then packed in ice and stored for the age of the solar system. Stardust has found fabulous things from the inner solar system, from extrasolar sources and even perhaps busted-up Pluto-like objects, but the sample is limited." JAXA is planning a direct sampling of a comet nucleus itself.

Earth's moon is another place to do some cosmoarchaeology. It has long been the Rosetta stone for understanding the impact history of the early solar system, connecting the relative ages provided by crater counts with the absolute dating of Apollo and Russian Luna samples. But the landers of the 1960s visited a limited range of terrains. They did not reach the Aitken basin, a continent-size crater on the far side, whose age may indicate when planet formation truly ended. NASA is now considering a robot to bring back a sample from there. It could run about half a billion dollars.

One oddity of the solar system is that the asteroids of the main asteroid belt apparently

CLEMENTINE SCIENCE GROUP, LUNAR AND PLANETARY INSTITUTE (*top*); JAXA (*middle* and *bottom*)

© 2007 SCIENTIFIC AMERICAN, INC.

formed before Mars, which in turn formed before Earth—suggesting that a wave of planet formation swept inward, perhaps instigated by Jupiter. But does Venus fit the progression? "There's no information," says planet formation expert Doug Lin of the University of California, Santa Cruz. "There's just *no* information." Between its acidic clouds, oceanic pressures and oven-cleaning temperatures, Venus is not exactly the friendliest environment for a lander. An NRC panel in 2002 recommended sending a balloon, which could touch down just long enough to collect samples and then repair to a cooler altitude to analyze them or forward them to Earth. The Soviet Union sent balloons to Venus in the mid-1980s, and the Russian space agency—which otherwise has fallen off the face of the earth when it comes to planetary exploration—now has plans for a new lander.

Studies of the origin of the planets overlap quite a bit with studies of the origins of life. Jakosky puts it thus: "Venus sits at the inner edge of the habitable zone. Mars sits at the outer edge. Earth sits in the middle. And understanding the differences between those planets is central to asking about life beyond our solar system."

ACTION PLAN

- **Return samples from a comet nucleus, the moon and Venus**

Right Stuff?

Scientists have a wide range of attitudes toward human spaceflight. Some think it incompatible with, even inimical to, scientific goals. Others think the two not only compatible but essentially the same thing—for them, curiosity-driven science and because-it's-there exploration are two sides of the same exploratory urge. Others think that humans will eventually want to leave the planet, out of either desire or desperation, even if the time has not yet come.

Whatever their views, researchers agree on several basic points. First, although astronauts can conduct useful science in space and on the moon and Mars, the cost of sending people greatly outweighs the scientific benefit. That may change in the future, as robots reach their limits, but for now a human program must be decided on its other merits; it is not primarily a scientific project. NASA administrator Michael Griffin has said explicitly that the moon/Mars initiative is not about science, although science gains by piggybacking on it.

Second, the space agency needs to respect the firewall between robotic missions and human missions, because the goals of these two wings of the space program are, for now, so distinct. Third, government initiatives and private flight each have something to contribute. With the retirement of the shuttle and then the International Space Station, Earth orbit can increasingly be left to the private sector, freeing NASA and other agencies to stay at the cutting edge.

Finally, if the nations of the world do send astronauts into space, they should at least give the travelers something worthy and inspiring to do. For most researchers, the space station, at least in its present form, does not count. Mars does. The moon is still hotly debated. —***G.M.***

CRUNCHING THE NUMBER$

NASA's budget is **$16.8 billion,** about **0.6 percent** of the total federal budget. Three fifths goes to human spaceflight, a third to science (for the planetary probes as well as space telescopes to explore the broader universe) and the rest to aeronautics.

The agency projects the new moon shot will run about **$100 billion** over the next decade. The Apollo program cost about the same.

This money comes from phasing out the shuttle and space station. President George W. Bush has retracted his earlier promise of a few extra billions, forcing a **20 percent cut** in the science program. Numerous missions have been canceled or put off.

NASA administrator Michael Griffin estimates that if the agency's budget just keeps up with inflation, astronauts could land on Mars in the **late 2030s**.

5 Break Out of the Solar System

Two years ago the venerable Voyager space probes went through a funding scare. NASA, desperate for money, said it might have to shut them down. The ensuing public outcry kept them going. Nothing that human hands ever touched has gone as far as Voyager 1: as of press date, 103 astronomical units (AU)—that is, 103 times as far from the sun as Earth is—and picking up another 3.6 AU every year. In 2002 or 2004 (scientists disagree), it entered the mysterious multilayered boundary of the solar system, where outgoing solar particles and inflowing interstellar gases go mano a mano.

But Voyager was designed to study the outer planets, not interstellar space, and its plutonium batteries are running down. NASA has long had a mind to dispatch a dedicated probe, and an NRC report on solar physics argued in 2004 that the agency should start working toward that goal.

The spacecraft would measure the abundance of amino acids in interstellar particles to see how much of the solar system's complex organics came from beyond; look for antimatter particles that might have originated in miniature black holes or dark matter; figure out how the boundary screens out material, including cosmic rays,

SCI AM

A Visual Summary

For more pictures and movies related to this article, go to **www.SciAm.com/ontheweb**

GETTY IMAGES

© 2007 SCIENTIFIC AMERICAN, INC.

which may affect Earth's climate; and see whether nearby interstellar space has a magnetic field, which might play a crucial role in star formation. The probe could act as a miniature space telescope, making cosmological observations unhindered by the solar system's dust. It might investigate the so-called Pioneer anomaly—an unexplained force acting on two other distant spacecraft, Pioneer 10 and 11—and pinpoint where the sun's gravity brings distant light rays to a sharp focus, as a test of Einstein's general theory of relativity. For good measure, scientists could aim the probe for a nearby star such as Epsilon Eridani, although it would take tens of thousands of years to get there.

Getting the thing hundreds of AU out within the lifetime of a researcher (and of a plutonium power source) would mean boosting it to a speed of 15 AU a year. The options boil down to large, medium and small—propelled, respectively, by an ion drive powered by a nuclear reactor, an ion drive powered by plutonium generators, or a solar sail.

The large (36,000-kilogram) and medium (1,000-kilogram) missions were honed in 2005 by teams led, respectively, by Thomas Zurbuchen of the University of Michigan at Ann Arbor and by Ralph McNutt of the Johns Hopkins University Applied Physics Laboratory. The small option seems the most likely to fly. ESA's Cosmic Vision program is now considering a proposal from an international team of scientists led by Robert Wimmer-Schweingruber of the University of Kiel in Germany. NASA might join in, too.

A solar sail 200 meters across could carry a 500-kilogram spacecraft. After launch from Earth, it would first swoop toward the sun, going as close as it dared—just inside Mercury's orbit—to get flung out by the intense sunlight. Like a windsurfer, the spacecraft would steer by leaning to one side or the other. Just before passing Jupiter's orbit, it would cast off the sail and glide outward. To get ready, engineers need to design a sufficiently lightweight sail and test it on less ambitious missions first.

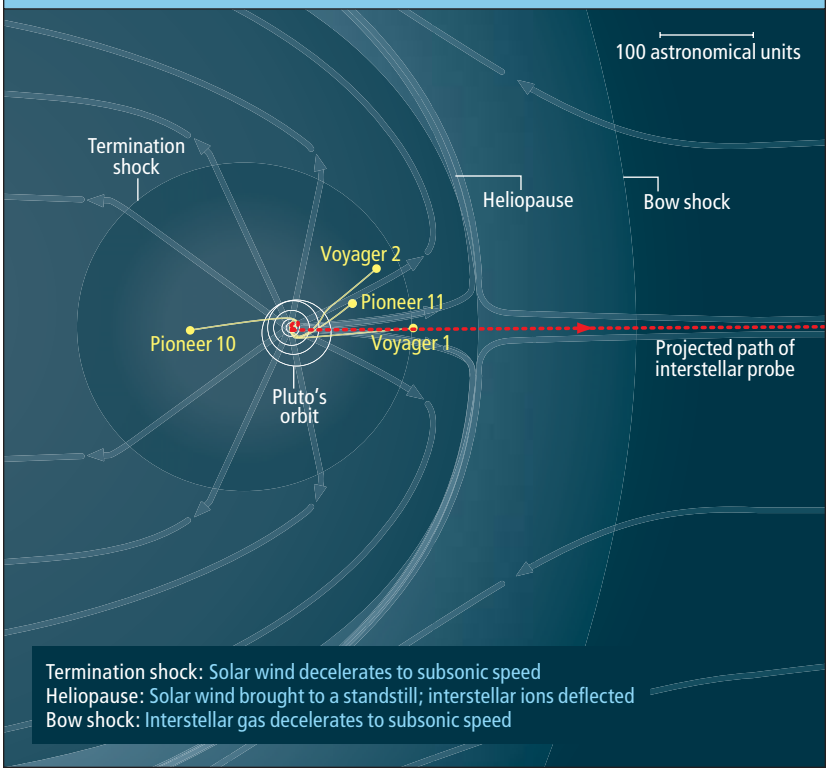
"Such a mission, be it ESA- or NASA-led, is the next logical step in our exploration of space," Wimmer-Schweingruber says. "After all, there is more to space than exploring our very, very local neighborhood." The estimated price tag is about $2 billion, including three decades' operating expenses. Studying the other planets has helped humans figure out how Earth plugs into a grander scheme, and studying our interstellar environs would do the same for the solar system at large.

ACTION PLAN

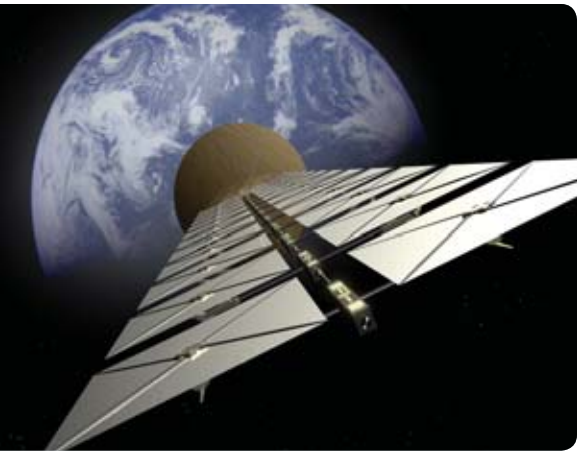
- **Begin developing and testing technology for an interstellar probe** ■

THE OUTER LIMITS

▼ An interstellar probe would explore the boundary region of the solar system, where gas flowing out from the sun pushes back the ambient interstellar gas. It would have the speed, endurance and instruments that the Pioneer and Voyager probes never did.



▼ **SOLAR SAIL is a big mirror (typically Mylar) that captures the momentum in sunlight. An interstellar probe needs one with a density of one gram per square meter, compared with 20 g/m² for current sails, but engineers think it doable.**



LUCY READING-IKKANDA (*top*); ESA (*bottom*)

MORE TO EXPLORE

Current positions of the Voyagers and Pioneers can be seen at **http://heavens-above.com/solar-escape.asp**

NASA's report on the asteroid threat is available at **http://neo.jpl.nasa.gov/neo/report2007.html**. For a critique, see **www.b612foundation.org/press/press.html**

NASA administrator Michael Griffin discusses the agency's future at **aviationweek.typepad.com/space/2007/03/human_space_exp.html**

National Research Council reports are available at **www.nap.edu/catalog/11937.html** (life on Mars), **11820.html** (Earth sciences), **11644.html** (science budget), **11135.html** (solar physics) and **10432.html** (solar system).

© 2007 SCIENTIFIC AMERICAN, INC.

How Does Consciousness Happen?

[THE AUTHOR]

Christof Koch is professor of cognitive and behavioral biology at the California Institute of Technology, where he teaches and has conducted research on the neuronal basis of visual attention and consciousness for more than two decades. He is an avid hiker and rock climber who has scaled several noted peaks.

■ **HIS THEORY: For each conscious experience, a unique set of neurons in particular brain regions fires in a specific manner.**

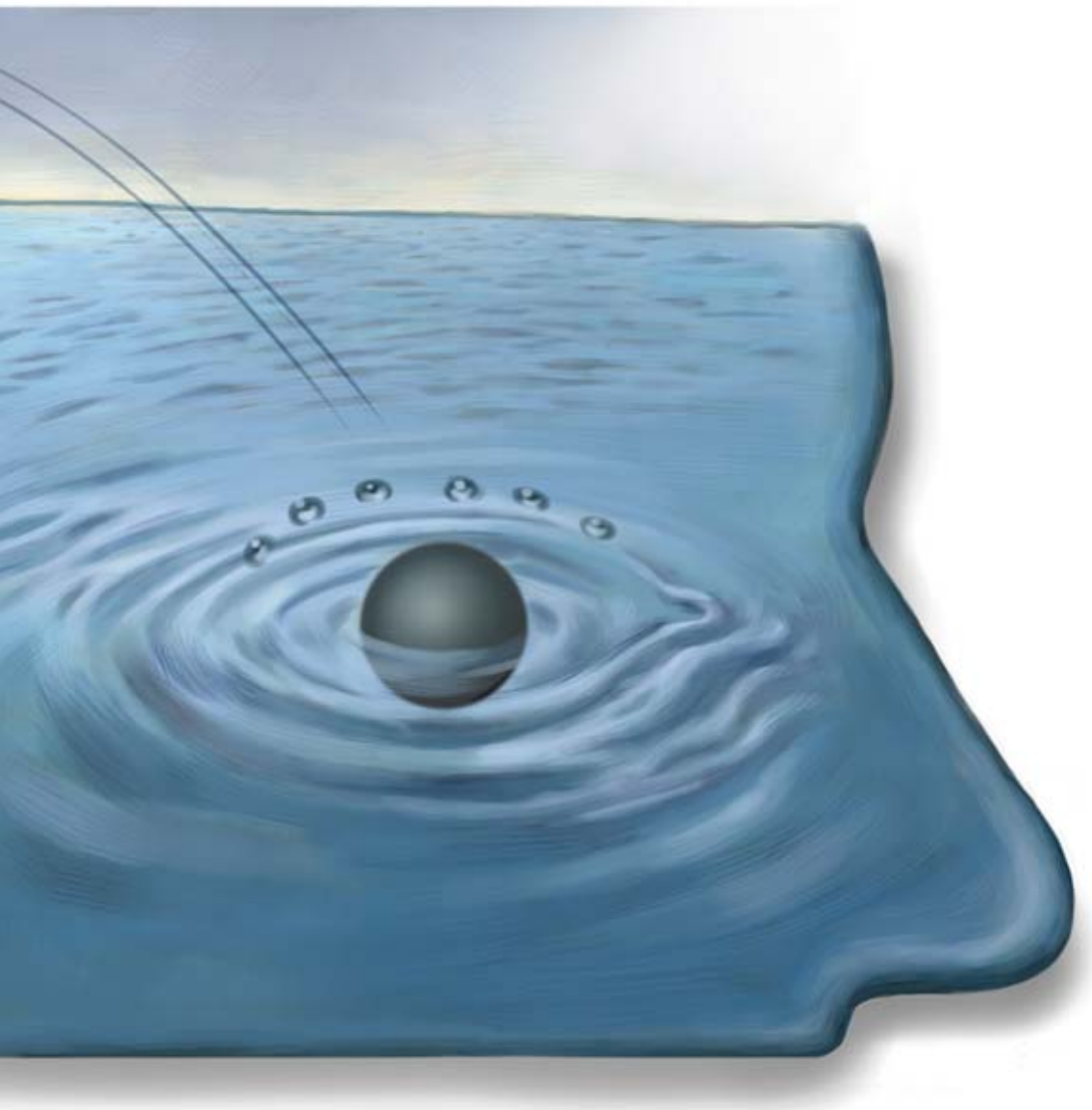
How brain processes translate to consciousness is one of the greatest unsolved questions in science. Although the scientific method can delineate events immediately after the big bang and uncover the biochemical nuts and bolts of the brain, it has utterly failed to satisfactorily explain how subjective experience is created.

As neuroscientists, both of us have made it our life's goal to try to solve this puzzle. We share many common views, including the important acknowledgment that there is not a single problem of consciousness. Rather, numerous phenomena must be explained—in particular, self-consciousness (the ability to examine one's own desires and thoughts), the content of consciousness (what you are actually conscious *of* at any moment), and how brain processes relate to consciousness and to nonconsciousness.

So where does the solution begin? Neuroscientists do not yet understand enough about the brain's inner workings to spell out exactly how consciousness *arises* from the electrical and chemical activity of neurons. Thus, the big first

JULIA BAIER (*Koch*); MATT COLLINS (*illustration*)

© 2007 SCIENTIFIC AMERICAN, INC.



Two leading neuroscientists, Christof Koch and Susan Greenfield, disagree about the activity that takes place in the brain during subjective experience

step is to determine the best neuronal correlates of consciousness (NCC)—the brain activity that matches up with specific conscious experiences. When you realize you are seeing a dog, what has happened among which neurons in your brain? When a feeling of sadness suddenly comes over you, what has happened in your brain? We are both trying to find the neuronal counterpart of each subjective experience that an individual may have. And this is where we differ.

Our disagreement over the best NCC emerged during a lively debate between us at the University of Oxford in the summer of 2006, sponsored by the Mind Science Foundation in San Antonio. Since then, we have continued to explore and challenge each other's views, a dialogue that has resulted in the article here. We are bound, nonetheless, by one fundamental commonality: our views stem primarily from neuroscience, not just philosophy. We both have considered a tremendous amount of neuroscientific, clinical and psychological data, and it is from these observations that our arguments arise.

—Christof Koch and Susan Greenfield

COURTESY OF SUSAN GREENFIELD

[THE AUTHOR]

Susan Greenfield is professor of pharmacology at the University of Oxford, director of the Royal Institution of Great Britain and member of the British Parliament's House of Lords. Her research focuses on novel brain mechanisms, including those underlying neurodegenerative diseases. Her favorite pastimes are squash and dancing.

■ **HER THEORY: For each conscious experience, neurons across the brain synchronize into coordinated assemblies, then disband.**

© 2007 SCIENTIFIC AMERICAN, INC.

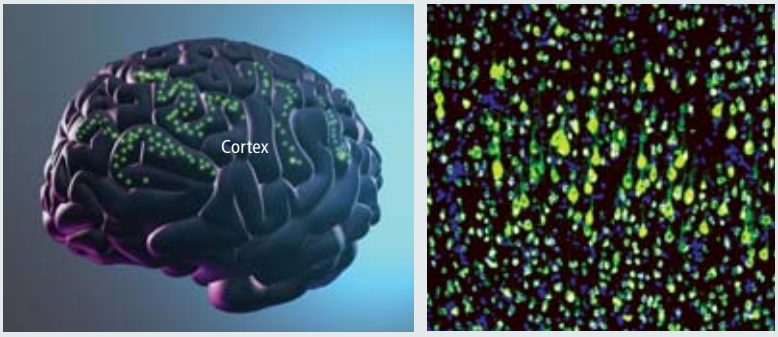
[BASIC ARGUMENTS]

CONSCIOUSNESS EXPLAINED

What happens in your brain when you see a dog, hear a voice, suddenly feel sad or have any other subjective experience?

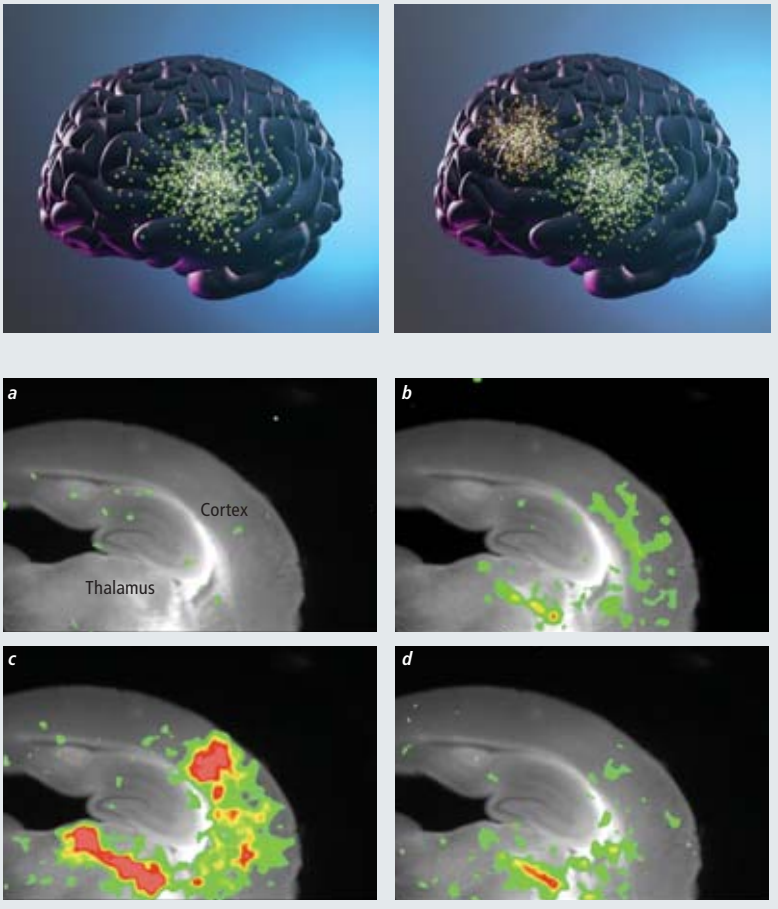
KOCH'S MODEL

A coalition of pyramidal neurons linking the back and front of the cortex fires in a unique way. Different coalitions activate to represent different stimuli from the senses (*left*). In a mouse cortex (*right*) these pyramidal cells (*green*) lie in brain layer 5, surrounded by nonneuronal cells (*blue*).



GREENFIELD'S MODEL

Neurons across the brain fire in synchrony (*green*) and prevail until a second stimulus prompts a different assembly to arise (*orange*). Various assemblies coalesce and disband moment to moment, while incorporating feedback from the body. In a rat brain (*bottom*), an assembly in the cortex forms (*a*, *b*), peaks (*c*), then decays (*d*) within 0.35 second after the thalamus is electrically stimulated.



Koch Speaks

"Specific groups of neurons mediate distinct conscious experiences."

Both Susan Greenfield and I are searching for the most appropriate neuronal correlates of consciousness. If we can find the right NCC, the direct cause-and-effect mechanisms that *create* consciousness may follow.

In my view, which has evolved since Francis Crick and I began investigating consciousness in 1988, every conscious percept (how the brain represents stimuli from the senses) is associated with a specific coalition of neurons acting in a specific way. There is a unique neuronal correlate of consciousness for seeing a red patch, another for seeing one's grandmother, a third for feeling angry. Perturbing or halting any neuronal correlate of consciousness will alter its associated percept or cause that percept to disappear.

Physiologically, the likely substrate for NCC is a coalition of pyramidal neurons—a type of neuron that communicates over long ranges—within the cerebral cortex. Perhaps only a million such neurons—out of the 50 billion to 100 billion in our heads—are needed to form one of these coalitions. When, say, Susan enters a crowded room and I see her face, a coalition of neurons suddenly chatters in concert for a fraction of a second or longer. The coalition reaches from the back of the cortex, where representations of visual stimuli are first processed, into the front of the cortex, which carries out executive functions such as providing perspective and enabling planning. Such a coalition would be reinforced if I paid attention to the stimulus of her image on my retina, which would strengthen the amplitude or the synchrony of the activity among the select neurons. The coalition sustains itself and suppresses competing coalitions by feeding excitatory signals back and forth among the neurons in the back and front of the cortex. If, suddenly, someone calls my name, a different coalition of neurons in the auditory cortex arises. This coalition establishes two-way communication with the front of the brain and focuses my consciousness on the voice, suppressing the earlier coalition representing Susan's face, which fades from my awareness.

One universal lesson from biology is that organisms evolve specific gadgets, and this is true for the brain. Nerve cells have developed myriad shapes and functions, along with specific wiring patterns among them. This heterogeneity is re-

STUART C. SEALFON AND POKMAN CHAN *Mount Sinai School of Medicine* (*mouse cortex*); MICHAEL HILL *University of Oxford* (*video stills*); ALFRED T. KAMAJIAN (*brains*)

© 2007 SCIENTIFIC AMERICAN, INC.

flected in the neurons that constitute NCC. It is here that I differ most from Susan. In my view, consciousness is not some holistic property of a large collection of firing neurons that are bathed in a solution of neurotransmitters, as she argues. Instead I maintain that specific groups of neurons mediate, or even generate, distinct conscious experiences.

And soon enough, the growing ability of neuroscientists to delicately manipulate populations of neurons will move us from observing that a particular conscious state is associated with some neuronal activity to pinpointing causation—observing that a given population is partially or wholly responsible for a conscious state.

But how do we determine which set of neurons, and what activity among them, constitutes a conscious percept? Do NCC involve all the pyramidal neurons present in the cerebral cortex at any given time? Or do they just involve a subset of long-range projection cells communicating between the frontal lobes and the sensory cortices in the back of the brain? Or do they involve neurons anywhere that are firing in synchrony?

Much of the contemporary work on NCC has concentrated on vision. Visual psychologists have perfected techniques to hide things from our conscious perception, like a magician who misdirects us so that we do not see what is happening in front of our eyes. One example is flash suppression, a phenomenon discovered by then graduate student Naotsugu Tsuchiya and myself in 2005. Perception of a small, stationary image shown to one eye—say, a faint, gray, angry face projected into the right eye—is completely suppressed by a stream of constantly changing color patches flashed into the other eye. This suppression can last for minutes, even though the scary face is perfectly visible if the viewer blinks his or her left eye; although legions of neurons in the primary visual cortex are firing vigorously in response to the stimulation from the left eye, they do not contribute to consciousness. This result is hard to explain in Susan's view that any coherent firing by a large collection of neurons is a correlate of consciousness. Researchers are using such illusions to find NCC in the brains of trained monkeys and humans.

Before Francis passed away, he and I offered several proposals about how consciousness works, based on experimental results. One is that NCC include pyramidal neurons that are strategically located in an output zone of the cerebral cortex known as layer 5. These cells send out signals to, and directly receive strong excit-

[POINT/COUNTERPOINT]

BRIAN CRONIN

Why does an alarm clock induce consciousness in a sleeping (unconscious) person?

Koch's view: Neurons in a region of the brain stem called the locus coeruleus respond to a sudden, large input from the auditory nerve. They spring into action, widely broadcasting a chemical signal to the thalamus and the cerebral cortex. Other neurons release the neurotransmitter acetylcholine throughout the brain. The net effect is that the cerebral cortex and its satellite structures become aroused. Once that occurs, a widespread but tightly interconnected grouping of neurons in the auditory cortex, and its counterparts in the front of the brain and in the medial temporal lobes that support planning and memory, establishes a stable coalition using recurrent feedback. This activity takes only a fraction of a second and causes you to become conscious of the alarm.

Greenfield's view: Any strong sensory stimulus, such as a bright light, will induce consciousness, so no one particular area of the brain can be responsible for waking you up. The alarm clock prompts consciousness not because of the quality of the stimulus (in this case, auditory) but because of its quantity (loudness). Transient neuronal assemblies—many neurons acting in concert—correlate with varying degrees of consciousness: the size of an assembly from one moment to the next is determined by how readily neurons can be corralled into transient synchrony. One key factor is the strength of sensory stimulation, the effects of which are akin to a stone thrown in a pond. The larger the stone, the more extensive the ripples on the water. The louder the alarm (or brighter the light), the more likely it will be to recruit an extensive assembly of neurons, and the more extensive the assembly, the more likely that you will be awakened.

© 2007 SCIENTIFIC AMERICAN, INC.

"Neuroscience needs a theory that can predict whether a fruit fly, a dog, an unresponsive Alzheimer's patient or the World Wide Web is conscious."

atory inputs from, another set of pyramidal neurons in a different region. Such an arrangement could implement a positive feedback loop, a coalition of neurons that, once triggered, would keep on firing until shut off by another coalition of neurons. These groups also fire over fractions of a second, much closer to the timescale of conscious awareness than single-neuron firings.

This notion about networks of neurons has received a boost from recent results by researchers at the Mount Sinai School of Medicine, Columbia University and the New York State Psychiatric Institute, working under Stuart C. Sealfon of Mount Sinai and Jay A. Gingrich of Columbia. Sealfon's and Gingrich's teams have demonstrated in genetically modified mice that hallucinogens—such as LSD, psilocybin (an ingredient of mushrooms) and mescaline—act on a type of molecule, called a serotonin receptor, found on the pyramidal cells that cluster in layer 5. The hypothesis that the mind-bending effects of hallucinogenic compounds come from activation of one receptor type on a specific set of neurons—rather than from "messing up" the brain's circuits in some holistic manner—can be further tested with molecular tools that can toggle layer 5 pyramidal cells on and off until the exact set of neurons being affected is identified.

A second proposal for how NCC underlie consciousness involves the claustrum, a sheet-like structure within the cortex. Remarkably the neurons composing this structure receive input from almost all regions of the cortex and project back to almost all as well. This structure may be perfectly situated to bind the activity of the sensory cortices into a single, coherent percept.

To advance these ideas, neuroscientists must sample the chattering electrical activity of a very large number of neurons at many locations. This work is delicate and difficult, but the miniaturization of electrodes is making it possible. Preliminary efforts confirm that specific groups of neurons express the types of perceptions that form our daily experiences.

None of these insights imply that one, 100 or even one million neurons living in a lab dish could be conscious. Neurons are part of vast networks and can generate consciousness only in that context. An analogy is helpful: although DNA molecules in a cell spell out the composition of the proteins in our bodies, many other molecules must also be present in the cell to construct and maintain those proteins.

The varying extent and provenance, or origin, of coalitions of neurons can also account for the different content of consciousness in infants, adults and animals. That any coalition can exist at all depends on the existence of arousal circuits in the brain stem and thalamus (which relays sensory inputs to the cortex) that are continuously active and that perfuse the cortex and its satellite structures with neurotransmitters and other substances. If a person's arousal circuits are silent—as they are when one is in deep sleep or under anesthesia or when one has suffered trauma akin to that of Terri Schiavo, the woman who fell into a persistent vegetative state that captivated the media—no stable coalition of cortical neurons can arise and the person is not conscious.

Although this model can be tested by physiological experiments, a valid criticism is that it is not a theory built from a set of principles—that is, it cannot predict what type of system has conscious experiences. Neuroscience needs a theory that predicts, based on physical measurements, which of the following organisms is conscious: a fruit fly, a dog, a human fetus five months after conception, an unresponsive Alzheimer's patient, the World Wide Web, and so on.

Some experts, including Giulio Tononi of the University of Wisconsin–Madison, are working on such theories. But we are still so ignorant about the brain that we can only speculate. Specific hypotheses that can be tested with today's technology will help. As Francis was fond of saying, what drove his and James Watson's 1953 discovery of the double-helical structure of DNA were experiments, not a theory of how genetic information might be encoded in molecules.

Fundamentally, my explanation is that qualitative, not quantitative, differences in neuronal activity give rise to consciousness. What matters is not the sheer number of neurons involved, as Susan stresses, but the informational complexity that they represent. A specific network of neurons is needed for a specific percept, not any random collection of neurons that become highly active. Furthermore, for full consciousness, a coalition of neurons must encompass both sensory representation at the back of the cortex as well as frontal structures involved in memory, planning and language. The brain works not by dint of its bulk properties but because neurons are wired up in amazingly specific and idiosyncratic patterns. These patterns reflect the accumulated information an organism has learned over its lifetime, as well as that of its ancestors, whose information is represented in genes. It is not crucial that a sufficient number of neurons are active together but that the right ones are active.

© 2007 SCIENTIFIC AMERICAN, INC.

Greenfield Speaks

"Consciousness is generated by a quantitative increase in the holistic functioning of the brain."

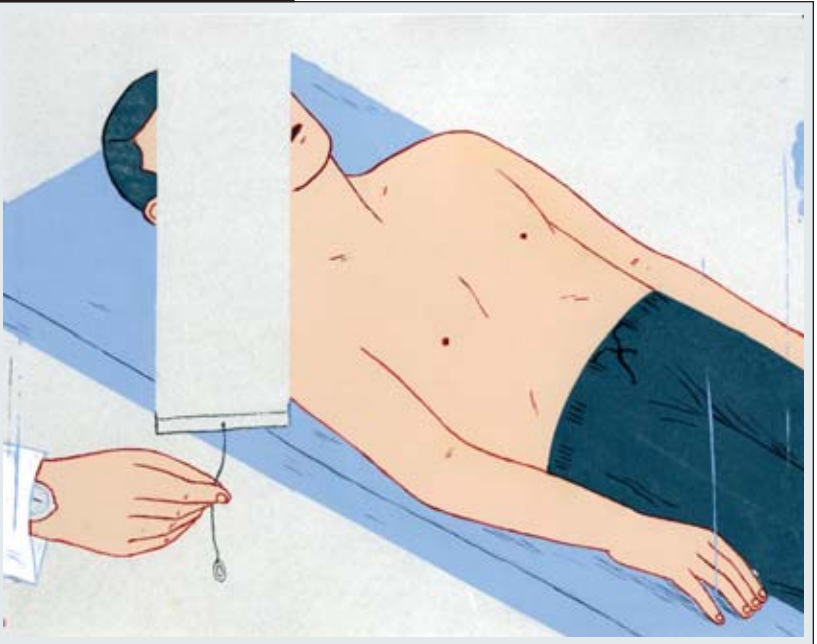
If neuronal correlates of consciousness are nothing more than the discharges of certain neurons and not others, as Christof Koch suggests, then consciousness resides in the neurons themselves. Yet Christof offers no explanation as to what qualitative property such neurons or regions have, compared with others. Moreover, if not even a million neurons can generate consciousness without being part of "vast networks," then the burden of identifying NCC shifts to describing what these networks are. By looking at specific brain connections for different forms of consciousness, Christof is guilty of a 21st-century form of phrenology, in which different functions are related directly to different brain regions, especially the cortex. His enthusiasm for the cortex should be tempered by the fact that many species, such as birds, have no cerebral cortex yet are considered conscious. Even if such compartmentalization were possible, it would not explain how consciousness is generated.

In my view, consciousness cannot be divvied up into different, parallel experiences. Indeed, we know that visual stimulation can change how we hear, and vice versa. This merging of the sensorium's components argues against concepts such as an isolated visual consciousness. Most important, either you are conscious or you are not. In Christof's lab, subjects are conscious throughout experiments performed on their neurons; therefore, it is not consciousness that the experiments manipulate but the *content* of that consciousness. Any consequent explanation is really a foray into answering "What is attention?" That question is certainly valid, but it is a different one from "What is consciousness?" I contend that to define the best NCC we must elucidate the difference between consciousness and unconsciousness.

My own starting assumption is that there is no intrinsic, magical quality in any particular brain region or set of neurons that accounts for consciousness. We need to identify a special *process* within the brain. And to be a truly robust correlate of consciousness, this neuronal process must account for a variety of everyday phenomena, including the efficacy of an alarm clock, the action of anesthetics, the distinction of dreams from wakefulness, the existence of self-consciousness, the possible difference between human and animal consciousness, and the possible existence of fetal consciousness. A more plausible view of consciousness is that it is not generated by a qualitatively distinct property of the brain but by a quantitative increase in the holistic functioning of the brain. Consciousness grows as brains grow.

BRIAN CRONIN

[POINT/COUNTERPOINT]



How do anesthetics work?

Koch: Today's anesthesiologists administer a diverse collection of chemicals. Yet all abolish consciousness. Scientists used to believe that anesthetics interfered systemically with lipids in the cellular membranes of neurons. But we now know that the compounds interfere with various neuronal processes by binding to certain membrane proteins. There is no single unique mechanism that causes consciousness to stop functioning. Among the most important causes, however, is that anesthesia strengthens synaptic inhibition, or reduces synaptic excitation, in large regions of the brain. Activity is not fully shut down, but the ability of groups of neurons to form stable coalitions is severely compromised. When neurons that encompass the back and the front of the cerebral cortex cannot set up synchronized communication, consciousness becomes impossible.

Greenfield: Anesthetics do not switch off any single brain area; they depress neuronal activity in different regions across the whole brain. Anesthetics therefore achieve their effect by altering an emergent property of the holistic brain: neuronal assemblies. As anesthetics reduce the size of neuronal assemblies, they reduce the degree of consciousness until it is nonexistent. This scenario also explains the different stages of consciousness that can occur as anesthesia takes effect, such as hyperexcitability and delirium. I have suggested elsewhere that people who have brains with underfunctioning neuronal connections, and who hence have small assemblies, often exhibit strong emotions and lack of reason—just the types of states many patients exhibit while anesthesia is taking effect and their assemblies are shrinking.

© 2007 SCIENTIFIC AMERICAN, INC.

[POINT/COUNTERPOINT]

Why is there a subjective difference between dreaming and wakefulness?

Koch: Although the brain is highly active during the rapid eye movement phase of sleep that is most associated with vivid dreams, the regional pattern of brain activity is quite distinct from that of wakefulness. In particular, the limbic system (loosely, the system of emotions and memory) is very active, but the parts of the frontal lobes that are involved in rational thought are subdued. In both dreaming and wakefulness, coalitions of neurons form, but they include neurons in different parts of the brain. During wakefulness, the coalitions include many more neurons in the prefrontal cortex, where reason and sensible narratives are imposed to order perceptions, but that activity is notably lacking during dreams. These features reflect the often bizarre and strong emotional content of dreaming.

Greenfield: Dreams most likely correlate with assemblies of neurons that are much smaller than those occurring when we are awake. The assemblies would be limited because no strong external stimuli are engaging large numbers of neurons. The transient recruitment of neurons during dreams is thus driven purely by response to spontaneous, intrinsic brain activity. And because the assemblies are not triggered by a sequential narrative of events in the outside world, the linkages among assemblies are haphazard, idiosyncratic or nonexistent, leaving dreams as random images or events. The lack of extensive, operational neuronal connections would also account for the notable absence of the checks and balances that normally characterize adult cognition when awake.

But what is the key neuronal mechanism in this process? The attempt to show a process-related correlate of consciousness has been inspired by various findings, including those of German neurophysiologist Wolf Singer. Singer demonstrated that a huge population of neurons between the thalamus and the cerebral cortex transiently fire together at a frequency of 40 times a second. But because the same activity can arise in this tissue kept alive in a lab dish, an additional condition must be a prerequisite for consciousness.

Neuroscientist Rodolfo Llinas of New York University Medical Center more recently suggested that this coordinated, transient firing sets up two complementary loops between the thalamus and the cerebral cortex that work in conjunction to maintain consciousness: a "specific" system relating to the content of consciousness and a "nonspecific" system relating to the arousal and alertness of consciousness. This account does indeed provide an explanation for why the strong sensory input of an alarm clock triggers full consciousness. Moreover, Llinas's model distinguishes between the consciousness of dreams and that of wakefulness; in dreams, there is no sensory input to feed the arousal loop, so only the content loop functions.

The central problem is that models developed by Llinas and others conceive of consciousness as an all-or-nothing condition. They fail to describe how the physical brain can accommodate the ebb and flow of a continuously variable conscious state. I favor an alternative. For more than a decade, scientists have known that the activity of tens of millions of neurons can synchronize for a few hundred milliseconds, then disband in less than a second. These "assemblies" of coordinating cells can vary continuously in just the right space and timescales for the here-and-now experience of consciousness. Wide-ranging networks of neurons assemble, disassemble and reassemble in coalitions that are unique to each moment. My model is that consciousness varies in degree from one moment to the next and that the number of neurons active within an assembly correlates with the degree of consciousness present at any given time.

This neuronal correlate of consciousness—the transient assembly—satisfies all the items on the shopping list of phenomena above. The efficacy of an alarm clock is explained as a very vigorous sensory input that triggers a large, synchronous assembly. Dreams and wakefulness differ because dreams result from a small assem-

BRIAN CRONIN

© 2007 SCIENTIFIC AMERICAN, INC.

bly driven by weak internal stimuli, whereas wakefulness results from a larger assembly driven by stronger external stimuli. Anesthetics restrict the size of assemblies, thus inducing unconsciousness. Self-consciousness can arise only in a brain large and interconnected enough to devise extensive neuronal networks. The degree of consciousness in an animal or a human fetus depends on the sizes of their assemblies, too.

Recall that neither Christof nor I is attempting to explain *how* consciousness arises. We are not attempting to answer what Australian philosopher David Chalmers has dubbed the "hard problem": determining how physiological events in the brain translate into what *you* experience as consciousness. We are seeking a correlation—a way to show how brain phenomena and subjective experiences match up, without identifying the all-important middle step of how a phenomenon causes an experience. Neuronal assemblies do not "create" consciousness but rather are indices of degrees of consciousness. Because an assembly's size and the corresponding degree of consciousness result from a variety of physiological factors—such as degree of connectivity, strength of stimuli and competition from other assemblies—each factor could eventually be manipulated experimentally. The assembly model's ability to generate falsifiable hypotheses *and* account for the diverse range of phenomena related to consciousness surely makes it particularly powerful.

An obvious criticism of the assembly model, which Christof articulated during our Oxford debate, is that it merely posits that "size is everything." But most of science is indeed "all about measurement"—the objective quantification of observations. Size *is* everything in science. Other skeptics say that assemblies are too vague a notion, but several researchers have revealed detailed characterizations of neuronal mechanisms that underlie the generation of assemblies lasting less than a second, such as Amiram Grinvald of the Weizmann Institute of Science in Rehovot, Israel, Ole Paulsen of Oxford and John G. Jefferys of the University of Birmingham in England.

Decisive tests in humans must await better noninvasive imaging techniques that have a time resolution commensurate with the milliseconds-long timescale of the formation and disbanding of neuronal assemblies. Once these techniques are available, we should be able to observe assemblies that correlate with the subjective experiences of, for example, neuropathic pain, depression and schizophrenia. Nevertheless, researchers have already observed the assembly model in action. In 2006 Toby Collins and others in my group at Oxford showed that in rats, the formation, activity and duration of assemblies correlate selectively with the action of anesthetics. Pilot observations in our laboratory, yet unpublished, also show that the number of neurons active in assemblies in the sensory cortex of an anesthetized rat reflects degrees of anesthesia. Earlier this year another member of my team, Subhojit Chakraborty, demonstrated that in rats, assemblies in the visual and auditory systems might serve as a good basis for distinguishing the subjectivity of seeing versus hearing.

Other criticisms relate to time and space. In epilepsy, for example, a prolonged neuronal assembly sustains a seizure, which equates with a *loss* of consciousness. But the whole point of assemblies as the appropriate NCC is that they are highly transient; a seizure acts as a jamming mechanism that prevents that transience, thus allowing a single assembly to last orders of magnitude longer than normal. Collins, Michael Hill, Eleanor Dommett and I have similarly suggested in a recent paper that anesthetics also may act as a jamming mechanism.

Another area of objection is that the assembly model does not have any spatial properties; there is no identified anatomical locus. But all too often we place far too much significance on localization as an end in itself. There is no need for a "center" for any given brain function, much less for consciousness.

A more plausible scenario would be that many different brain regions, in generating highly transient assemblies, converge as inputs to a space-time manifold. The present difficulty is that we cannot describe such a manifold using current experimental techniques. Perhaps the manifold could eventually be modeled mathematically. Such models and their interactions may be the way forward.

A final problem, and one that applies to NCC at the basic level, is how they might be harnessed to tackle the hard problem: determining how physiological events in the brain translate into what you experience as consciousness. We will not be in a position to find a solution until we know what kind of evidence would satisfy us: A brain scan, a performing rat, a robot, a formula? Or perhaps an induced change in one's subjective state, such as if Christof's brain could be manipulated so that he would experience the world as I do—and even agree with me. ■

"There is no need for a 'center' for any given brain function, much less for consciousness."

MORE TO EXPLORE

The Private Life of the Brain. Susan Greenfield. John Wiley & Sons, 2000.

A Framework for Consciousness. Francis Crick and Christof Koch in *Nature Neuroscience,* Vol. 6, pages 119–126; February 2003.

The Quest for Consciousness: A Neurobiological Approach. Christof Koch. Roberts & Company Publishers, 2004.

A Neuroscientific Approach to Consciousness. Susan A. Greenfield and T.F.T. Collins in *Progress in Brain Research,* Vol. 150, pages 11–23; 2005.

The authors would like to thank Joseph Dial of the Mind Science Foundation for his support. Susan Greenfield thanks Michael Hill, Nicholas Shea and Kathleen Taylor for their insights.

© 2007 SCIENTIFIC AMERICAN, INC.

THE DIAMOND AGE OF SPINTRONICS

Quantum electronic devices that harness the spins of electrons might one day enable room-temperature quantum computers—made of diamond

By David D. Awschalom, Ryan Epstein and Ronald Hanson

KEY CONCEPTS

- Electrons carry both charge and spin, but only spintronic devices exploit the two properties simultaneously to achieve innovative capabilities.
- Spintronics brings us disk-drive read heads and nonvolatile memory chips today and perhaps instant-on computers with reconfigurable chips tomorrow.
- Synthetic semiconducting diamond may be the new silicon for a future era of quantum spintronic technology that manipulates single spins, enabling quantum computers and other quantum information devices.

—The Editors

Diamond has a track record of extremes, including ultrahardness, higher thermal conductivity than any other solid material and transparency to ultraviolet light. In addition, diamond has recently become much more attractive for solid-state electronics, with the development of techniques to grow high-purity, single-crystal synthetic diamonds and insert suitable impurities into them (doping). Pure diamond is an electrical insulator, but doped, it can become a semiconductor with exceptional properties. It could be used for detecting ultraviolet light, ultraviolet light-emitting diodes and optics, and high-power microwave electronics. But the application that has many researchers excited is quantum spintronics, which could lead to a practical quantum computer—capable of feats believed impossible for regular computers—and ultrasecure communication.

Spintronics is an advanced form of electronics that harnesses not just the electrical charge of electrons (as in conventional electronics) but also a property called spin that makes electrons act like tiny bar magnets. Your computer probably already contains the first and most rudimentary commercial application of spintronics: since 1998 hard-drive read heads have used a spintronic effect called giant magnetoresistance to detect the microscopic magnetic domains on a disk that represent the 1s and 0s of the data it contains.

Another spintronic device, one that you may find in new computers in the next few years, is magnetoresistive random-access memory (MRAM). As with a hard drive, MRAM stores information as magnetization and therefore is nonvolatile, meaning that the data are not lost when the device's power is turned off. The readout is done electrically, just like any other charge-based memories [see "Spintronics," by David D. Awschalom, Michael E. Flatté and Nitin Samarth; Scientific American, June 2002]. Freescale Semiconductor, a spin-off of Motorola, began selling the first MRAM in 2006.

Nonvolatile memory chips could lead to computers that will not need to reload programs laboriously from a hard drive every time they are switched on. Instead a computer would be ready within a fraction of a second to proceed from

KENN BROWN

© 2007 SCIENTIFIC AMERICAN, INC.

where it left off (much like handhelds today) because all the necessary programming and data would remain ready and waiting in the chip.

More advanced spintronic technologies that are in the early research stages—such as spin transistors, which would make use of spin in controlling current flow—could enable computer chips with logic circuitry capable of being reconfigured on the fly.

Quantum Electronics

Devices such as read heads and MRAM chips represent one class of spintronics, in which the spins of large numbers of electrons are aligned the same way, as with a collection of toy tops all spinning clockwise on the floor. These so-called spin-polarized electrons typically flow through some part of the device, forming a spin-polarized current, or spin current, that is highly analogous to a polarized beam of light. Researchers have made many exciting advances in this area in the past few years, including discovery of ways to generate and manipulate spin polarization in semiconductors without relying on a magnetic material or relatively bulky wiring to generate a magnetic field. In particular, our group and others have observed a potentially very useful phenomenon called the spin Hall effect [see "Spin Control for the Masses," on page 88].

Much further from store shelves is the second class—quantum spintronics—which involves the manipulation of individual electrons to exploit the quantum properties of spin. Quantum spintronics could provide a practical way to carry out quantum information processing, which replaces the definite 0s and 1s of ordinary computing with quantum bits, or qubits, capable of being 0 and 1 simultaneously, a condition called a quantum superposition [see "Rules for a Complex Quantum World," by Michael A. Nielsen; Scientific American, November 2002].

Quantum computers, if they can be built, would exploit superpositions of qubits to perform a kind of parallel processing that would be extremely effective for certain tasks, such as searching databases and factoring large numbers. Efficient number factoring looms large because it would render obsolete cryptographic

© 2007 SCIENTIFIC AMERICAN, INC.

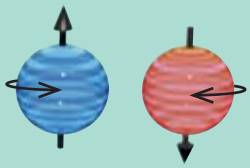
[THE BASICS]

SPIN AND ITS USES

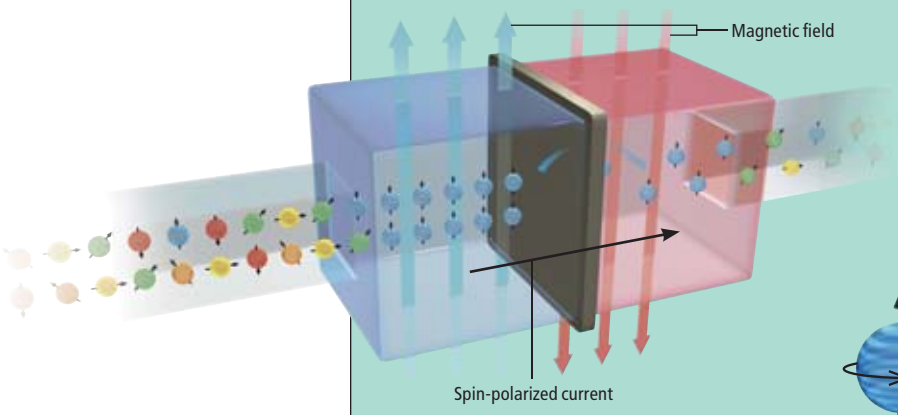
WHAT IS SPIN?

In addition to their mass and electric charge, electrons have an intrinsic quantity of angular momentum called spin, almost as if they were tiny spinning balls.

Scientists represent spin with a vector. For a sphere spinning "west to east," the vector points "north," or "up." It points "down" for the opposite spin.



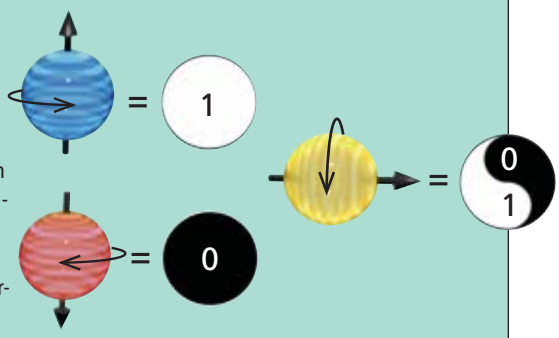
Magnetic tunnel junction



TWO KINDS OF SPINTRONICS

The first class of spintronic devices uses spin-polarized electric currents, in which the electrons have their spins aligned. The earliest of these devices, such as magnetic tunnel junctions (*left*), rely on magnetic fields to polarize the electrons and are already commercially available.

The second class controls individual electrons, using them to represent quantum bits (qubits) and to carry out quantum information processing. If spin "up" is a 1 and spin "down" a 0, a tilted electron spin is a quantum superposition of 0 and 1. These devices all remain highly experimental and include the diamond-based spintronics.



codes that are widely used, including for secure communication over the Internet. Anyone with a large enough working quantum computer (say, an intelligence or law enforcement agency or a corporation) would be able to decode countless formerly secret messages at will.

Perhaps the greatest impact of a future quantum computer will lie in its unique capability to simulate, or model, other quantum systems, a task that current computers are hopelessly bad at. For example, quantum simulations will be required to understand the behavior of matter at the nanometer scale and could therefore bring huge advances in physics, chemistry, materials science and biology.

This exciting prospect has led to a worldwide race to find the most suitable system for storing and processing quantum information. The most advanced quantum information-processing units to date are arguably spins of ions trapped in electromagnetic fields. But these systems have the disadvantages of requiring an ultrahigh vacuum and complex trapping architectures to hold the individual particles in place and isolated from disturbances. Developing chips with large numbers of such traps on them is a major challenge. In contrast, solid-state qubits, which reside directly in a solid substrate, could allow developers to build on decades of experience fabricating semiconductor chips.

Yet many questions have loomed large for researchers hoping to implement solid-state quantum computing: Can spins in solids be individually addressed and controlled? Can scientists come up with suitable interactions to implement quantum logic gates reliably? Can spins in solids maintain quantum information long enough to perform a useful number of operations on that information? In the past few years, all these questions have been answered positively. Surprising-

ILLUSTRATIONS BY JEAN-FRANCOIS PODEVIN

© 2007 SCIENTIFIC AMERICAN, INC.

ly, one of the most promising host materials for spins turned out to be diamond.

Glitter of Diamonds

The diamond we use in our experiments looks very different from the sparkling gemstones found in jewelry. Recent advances in materials science make it possible to synthesize thin films of diamond—typically a few hundred nanometers thick over areas as large as many square centimeters—by chemical vapor deposition. In this process, a gas made up of carbon-containing molecules (often methane) and hydrogen is broken down into individual atoms (for example, by high-power microwave radiation), allowing the carbon atoms to deposit on a silicon substrate. The diamond that forms can be extremely pure but often consists of many small crystals, or grains, with grain sizes ranging from nanometers to microns depending on the conditions in the chamber. The best device performance usually comes from using single-crystal diamond, in which diamond's characteristic tetrahedral lattice of carbon atoms is uninterrupted by the disorderly grain boundaries, which degrade the quality of the material for both optics and electronics. The ability to engineer diamond into many forms will likely have a profound effect on electronics, both conventional and quantum.

A key property of diamond for quantum elec-

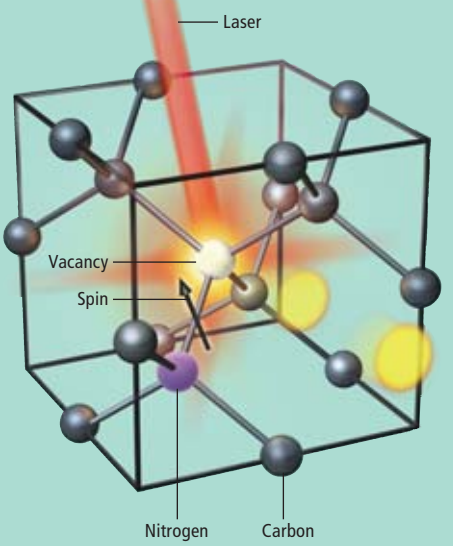
[DIAMOND SPINTRONICS]

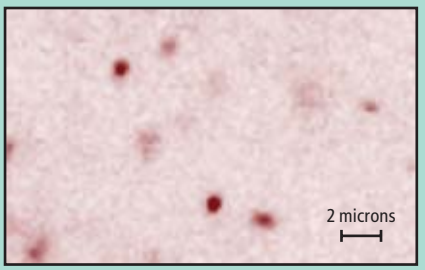
A MAGICAL IMPURITY

As with semiconductors in conventional electronics, the key to making diamond functional for quantum spintronics is doping it with an impurity, in this case a so-called nitrogen-vacancy (N-V) center.

At an N-V center, two adjacent sites in diamond's tetrahedral lattice of carbon atoms are altered. One has a nitrogen atom instead of a carbon, and the other has an empty space. Electrons orbit in the vacancy and around the adjacent four atoms and carry a spin that quantum applications can exploit.

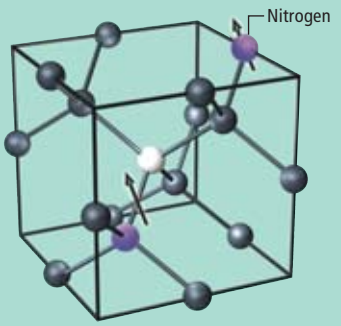
For example, a laser can repeatedly excite an electron at the N-V center, which each time emits a single photon in a specific quantum state when it decays back to its unexcited state. Researchers have used diamond in this way to demonstrate quantum cryptography prototypes, which rely on a steady supply of single photons.





N-V centers in diamond show up as bright spots (*red*) when pumped by a laser. Centers whose spin is in state 1 are much brighter than centers whose spin is in state 0. Radio-frequency waves tuned to a precise frequency change the N-V centers back and forth between 0 and 1, passing through transitional states that are quantum superpositions of the two.

Inserting a second nitrogen atom near the N-V center provides a system of two coupled qubits that enables logic processing. The frequency required to flip the N-V center's qubit is now slightly lower or higher, depending on the state of the second nitrogen. Applying waves at the higher frequency can therefore flip the N-V qubit only if the other qubit is 1. That operation is known as a controlled NOT logic gate, which enables arbitrary quantum computations.



ENABLED BY SPINTRONICS

Very high densities of data storage on disk drives.

Nonvolatile memory chips.

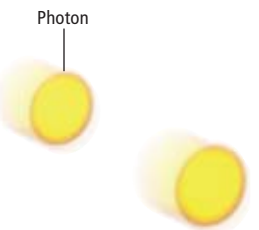
"Instant-on" computers.

Chips that both store and process data.

Chips operating at higher speeds and consuming less power than conventional ones.

Chips with logic gates that can be reconfigured on the fly.

Quantum cryptography and quantum computing at room temperature.



ILLUSTRATIONS BY JEAN-FRANCOIS PODEVIN; DATA BY DAVID D. AWSCHALOM

© 2007 SCIENTIFIC AMERICAN, INC.

[SPIN CURRENTS]

SPIN CONTROL FOR THE MASSES

By Yuichiro K. Kato

Spintronic devices exploit spin, a property of electrons that makes them like tiny bar magnets. There are two classes of such devices—those that manipulate the spins of single electrons [*see main text*] and those that control large groups of spin-polarized electrons flowing en masse in semiconductors (spin currents). Along with working toward single-electron devices, researchers are making exciting discoveries in controlling spin currents.

I was fortunate enough to play a role in these advances while I was a graduate student working in David D. Awschalom's group at the University of California, Santa Barbara, from 2000 to 2005. In particular, we found new ways to generate and manipulate spin polarization. We also observed for the first time a phenomenon called the spin Hall effect, which may provide a way to sort and route electrons based on the direction of their spins.

Because spins behave like tiny magnets, people control them by applying magnetic fields. Producing magnetic fields usually requires magnetic materials or external magnets, however. Instead, using electrical fields might enable smaller, faster spintronic devices that are simpler to fabricate because electric fields are easier to confine to small regions and easier to produce with high frequencies (which enable faster operations). Unfortunately, spins, like all magnets, do not respond to electric fields under normal circumstances.

A relativistic effect comes to the rescue: electrons that move perpendicular to an electric field "see" a weak magnetic field mixed in with the electric one. The magnetic field influences the electron's spin. This interaction is called spin-orbit coupling because physicists first studied it in relation to electrons "orbiting" in the electric field of atomic nuclei.

The Santa Barbara group initially studied this effect in gallium arsenide, a semiconductor commonly used in electronics. We saw that when we moved packets of spin-polarized electrons through this material, the spins rotated as if they were in a magnetic field. The phantom magnetic field could also align the spins of unpolarized electrons.

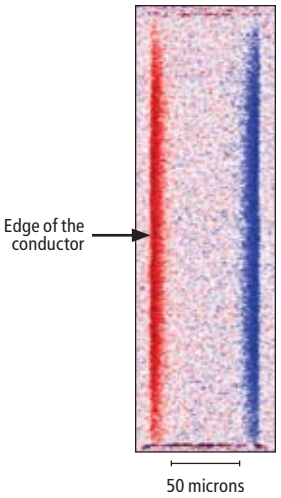
Spin-orbit coupling also gives rise to the spin Hall effect, which was predicted in 1971 by Michel D'yakonov and Vladimir Perel of the Ioffe Institute in Leningrad. It is named by analogy with the Hall effect (discovered in 1879 by Edwin Hall), in which opposite charges build up on each side of a material that carries a current in a magnetic field (*top right*). In the spin Hall effect, a small spin polarization accumulates on the edges of a material carrying an electric current (*bottom right*), but without requiring a magnetic field. This effect would be another nonmagnetic way to generate spin polarization and to direct electrons according to their spin orientation.

In late 2004 Robert C. Myers (another graduate student), Arthur C. Gossard, Awschalom and I reported seeing the expected spin polarization at the edges of a slab of gallium arsenide chilled to 30 kelvins. A few months later a group led by Jörg Wunderlich at Hitachi Laboratory in Cambridge, England, published observations of the spin Hall effect involving holes (absences of electrons). About a year ago the Awschalom group went on to demonstrate the spin Hall effect at room temperature in the semiconductor zinc selenide.

Taken together, these discoveries offer exciting possibilities for developing spin-based semiconductor technology.

Yuichiro K. Kato is associate professor in the Institute of Engineering Innovation at the University of Tokyo.

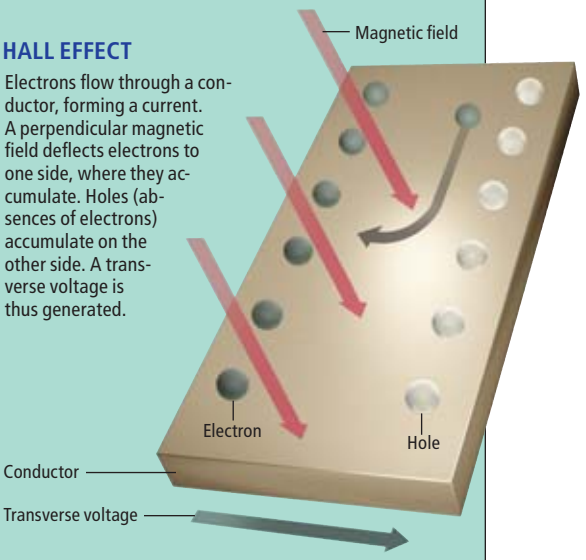
SPIN HALL EFFECT EXPERIMENTALLY OBSERVED



MEASUREMENT made in 2005 detected electrons with opposite spin polarizations (*red, blue*) accumulating at the edges of a conductor that had a current passing through it lengthwise.

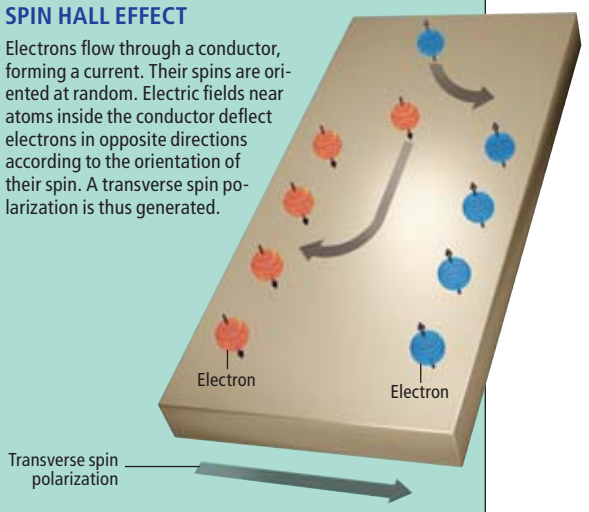
HALL EFFECT

Electrons flow through a conductor, forming a current. A perpendicular magnetic field deflects electrons to one side, where they accumulate. Holes (absences of electrons) accumulate on the other side. A transverse voltage is thus generated.



SPIN HALL EFFECT

Electrons flow through a conductor, forming a current. Their spins are oriented at random. Electric fields near atoms inside the conductor deflect electrons in opposite directions according to the orientation of their spin. A transverse spin polarization is thus generated.



ILLUSTRATIONS BY JEAN-FRANCOIS PODEVIN; DATA BY DAVID D. AWSCHALOM

© 2007 SCIENTIFIC AMERICAN, INC.

tronics is the large amount of energy needed to dislodge an electron so that it can flow through the material. Physicists visualize the states that electrons can have in a solid as bands of different energy forming a ladder of unevenly spaced rungs. For semiconductors, the two important bands are the valence band, which is the highest band containing electrons, and the empty conduction band just above it, in which electrons can flow freely. The size of the energy gap, or band gap, between these two bands in diamond is 5.5 electron volts, about twice as much energy as present in a visible-light photon and five times as large as the band gap in silicon.

Generally electrons in a semiconductor cannot have an energy that lies in the gap, but impurity atoms added to the material can introduce discrete states in the gap, like additional thin rungs to the ladder. Diamond's gap is big enough that two of these states can differ by an energy as large as that of a visible-light photon. Thus, optical-wavelength light can excite an electron at an impurity atom from one discrete state to another without knocking it all the way to the conduction band. When the electron falls back into its lower energy state, it emits a photon with a frequency corresponding to the energy-level difference—the process commonly known as fluorescence. Under continuous illumination, the optical excitation and relaxation process repeats over and over, and an impurity can emit millions of photons per second. In 1997 a group led by Jörg Wrachtrup, who was then at the University of Technology in Chemnitz, Germany, detected individual impurities in diamond fluorescing in this way, igniting a wave of research in the optical detection of single impurities.

The particular impurity that Wrachtrup's group detected in those first experiments consisted of a nitrogen atom in place of one carbon atom and an adjacent void where another carbon usually would be, which is known as a nitrogen-vacancy (N-V) center. The N-V center in diamond has a number of remarkable properties that make it the favorite subject of research for many groups around the world. Interestingly, the void plays a crucial role: the N-V center is quite different from a single nitrogen atom without a neighboring vacancy. The electrons in the N-V center move in orbits that span the vacancy and its three neighboring carbons and spend only a little time near the nitrogen. Because of these molecularlike orbits, it is convenient to think of the N-V center as being a single impurity rather than a somewhat odd composite of a nitrogen atom and a vacancy.

Single impurities, such as an N-V center, emit one photon at a time—a vital property for the burgeoning field of quantum cryptography [see "Best-Kept Secrets," by Gary Stix; Scientific American, January 2005]. Quantum cryptography systems transmit information in the form of single photons carrying one qubit apiece. The laws of physics guarantee that an eavesdropper cannot intercept the photons without disturbing the qubits in ways that the intended recipient can detect. In 2002 Philippe Grangier and his coworkers at the Institute of Optics in Orsay, France, demonstrated the first quantum cryptography prototype system based on a pulsed source of single photons. This breakthrough relied on having an extremely stable and reliable single-photon source—an N-V center in diamond.

The N-V center electrons also carry a spin state, one which can be polarized conveniently with optical-wavelength light. And whereas other spin systems in the solid state typically must be cooled to very low temperatures to be polarized, the N-V center spin naturally goes into a specific spin state under optical illumination even at room temperature. Furthermore, experimenters soon discovered that one of the spin states fluoresces much more brightly than the others. Thus, fluorescence intensity can be used for spin-state readout—bright for state "1," dim for state "0."

Diamonds Are Forever

During the past few years, our group at the University of California, Santa Barbara, has developed a single-photon imaging technique to observe such individual spins and their orientation in the diamond lattice and to manipulate them. We have thereby studied how single spins interact with their environment—in this case the diamond that surrounds them—a topic of fundamental importance in developing quantum applications. The interactions of the N-V centers with nearby atoms have allowed us to observe so-called dark spins in diamond—impurity nitrogens without an associated vacancy that are invisible to optical detection on their own.

Crucially, as observed in these measurements, spins in diamond are extremely stable against environmental disturbances. Indeed, one of the most exciting aspects of the N-V center is that it exhibits quantum behavior even at room temperature. Quantum phenomena tend to be washed out by thermal excitations, and many solid-state quantum effects require extremely cold temperatures, making them hard to study and

DIAMOND'S MANY FACETS

The name "diamond" derives from the ancient Greek word *adamas,* meaning "invincible."

Diamond is the hardest known naturally occurring substance. The hardest known substance overall is aggregated diamond nanorods, which are 1.11 times as hard as diamond by one measure.

Diamond conducts heat better than any other solid material.

Diamond has a high refractive index (2.4, compared with about 1.5 for glass).

Pure diamond is an electrical insulator (blocking all current flow), but when doped with impurities it can become semiconducting.

Because the energy gap between bound electrons and conduction electrons is large in semiconducting diamond, it is transparent to ultraviolet light and so could be used for ultraviolet detectors and light-emitting diodes. High-power electronics is another application.

By allowing impurities to be excited without becoming ionized, the large energy gap is one of the keys to the quantum spintronics applications.

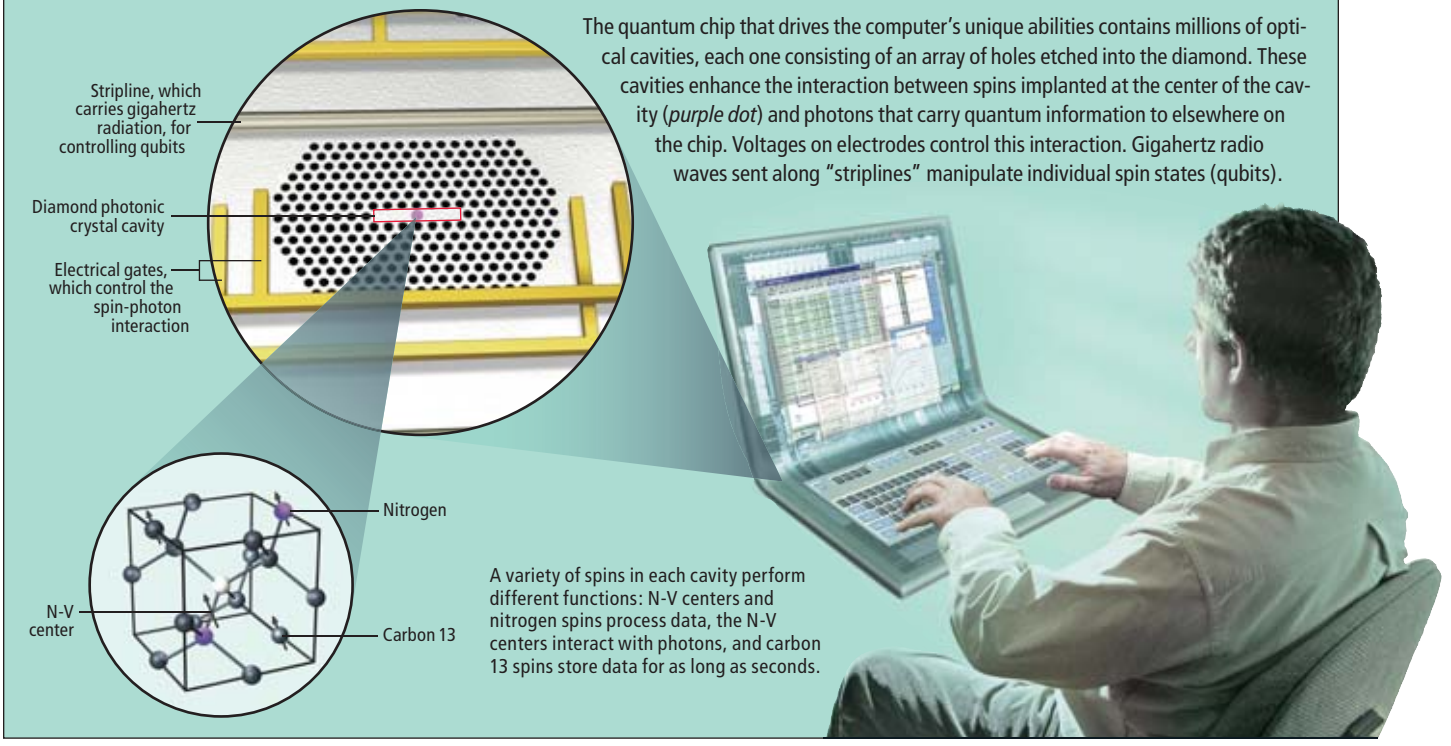
Quantum spin states of impurities in diamond can retain their quantum character for usefully long times (about a millisecond), even at room temperature.

© 2007 SCIENTIFIC AMERICAN, INC.

[QUANTUM FUTURE]

DIAMOND MICROPROCESSOR

In the future, people wishing to carry out certain specialized tasks may use quantum computers based on diamond spintronics.



The quantum chip that drives the computer's unique abilities contains millions of optical cavities, each one consisting of an array of holes etched into the diamond. These cavities enhance the interaction between spins implanted at the center of the cavity (*purple dot*) and photons that carry quantum information to elsewhere on the chip. Voltages on electrodes control this interaction. Gigahertz radio waves sent along "striplines" manipulate individual spin states (qubits).

A variety of spins in each cavity perform different functions: N-V centers and nitrogen spins process data, the N-V centers interact with photons, and carbon 13 spins store data for as long as seconds.

even harder to turn into practical technology.

In this regard, spins in solid materials typically suffer from two problems. The first is an interaction called spin-orbit coupling, which involves the electron's spin and its orbital motion. The second is magnetic interactions with other spins, such as the spins of the nuclei that make up the lattice. In diamond, both these effects are very weak. For example, the nuclei of carbon 12, which makes up 99 percent of natural carbon, have zero spin and thus no effect on the spin of an N-V center. Because it is so immune to outside disturbances of this kind, the quantum state of the N-V center spin can be used to encode quantum information even at room temperature.

Of course, "immune" is a relative term. The quantum information stored in an N-V center spin state is lost after about one millisecond in high-purity diamond at room temperature. This loss is equivalent to a bit being flipped in a regular computer. As with such errors in ordinary computers, mistakes in qubits can be corrected provided the error rate is low enough. A rule of thumb for quantum error correction is that at most one in 10,000 operations may fail; any more than that and the procedure becomes a losing battle, with the extra data and operations needed to perform the correction themselves introducing too many new errors.

How does the N-V center in diamond measure up against the one-in-10,000 criterion? Radio-frequency radiation guided to the N-V center through on-chip waveguides can deliberately change the N-V center spin within 10 nanoseconds. About 100,000 such operations can occur in the millisecond-long lifetime of the spin's quantum state, and thus the error rate will be very roughly one failure in 100,000 operations. This rate is well below the threshold and is better than any other system of solid-state qubits to date.

Quantum cryptography requires only a sequence of individual qubits, but for quantum computation the qubits must interact to produce new qubits, a process that is analogous to how logic gates process pairs of input bits to produce an output in ordinary computers. For example, an AND gate produces an output of 1 if both inputs are 1 and produces 0 otherwise. Quantum logic gates must do similar operations and must also accept quantum superpositions of bits as inputs and produce superpositions as outputs. The next step toward quantum information processing with impurity spins is controlling the coupling between two spins to perform quantum logic.

MORE TO EXPLORE

A Hall of Spin. Vanessa Sih, Yuichiro Kato and David D. Awschalom in *Physics World,* Vol. 18, pages 33–37; 2005.

Two Groups Observe the Spin Hall Effect in Semiconductors. Charles Day in *Physics Today,* Vol. 58, No. 2, pages 17–19; February 2005.

Challenges for Semiconductor Spintronics. David D. Awschalom and Michael E. Flatté in *Nature Physics,* Vol. 3, pages 153–159; 2007.

Spins in Few-Electron Quantum Dots. R. Hanson, L. P. Kouwenhoven, J. R. Petta, S. Tarucha and L.M.K. Vandersypen in *Reviews of Modern Physics* (in press).

ILLUSTRATION/PHOTOILLUSTRATION BY JEAN-FRANCOIS PODEVIN

© 2007 SCIENTIFIC AMERICAN, INC.

Our group and Wrachtrup's have studied an interaction that could carry out quantum logic by using two spins that are near each other in the diamond lattice. Specifically, we have measured how the spin on an N-V center interacts with another spin on a nearby nitrogen impurity (with no vacancy). The interaction is largely magnetic dipole coupling, essentially the same as the force that makes two macroscopic bar magnets align with north poles facing south poles.

The interaction works as follows. The 0 and 1 states of an N-V center have somewhat different energies. The energy difference, or splitting, between the 0 and the 1 is much smaller than the energy of an optical photon, and instead gigahertz radio waves will drive the spins back and forth between 0 and 1 and superpositions thereof. When the N-V center is close to another nitrogen atom, the splitting of its 0 and 1 states depends on the other nitrogen's spin state. This dependence makes possible a controlled NOT (CNOT) gate, in which one qubit is inverted if and only if the other qubit is a 1. The gate would work by using radio waves tuned to the frequency that will flip the N-V center provided the nitrogen spin is a 1. If the nitrogen spin is a 0, the N-V center's energy splitting will be different and the radio waves will not affect it.

The CNOT gate is quite special: we can compose any arbitrary quantum operation on any number of qubits by combining CNOT gates acting on pairs of qubits and rotations of individual qubits (which can also be carried out by applying radio waves to spins; individual spins could be addressed by bringing the radiation to them along special circuits called striplines). Demonstrations of a CNOT gate and qubit rotation are therefore major research goals.

Longer distance interactions between N-V spins in diamond may be possible by using photons as mediators. On-chip optical devices such as waveguides made of the same diamond substrate could route the photons. Integrating the N-V centers in structures called optical cavities, in which light forms standing waves, would enhance the strength of the interaction between the spins and the photons. At Santa Barbara, in a collaboration with Evelyn Hu and her students, we recently demonstrated proof-of-concept photonic crystal cavities. Each "optical cavity" consists of a region of diamond with a honeycomb of holes etched into it. The holes work to confine and amplify light at the center of the structure [see "Photonic Crystals: Semiconductors of Light," by Eli Yablonovitch; Scientific American, December 2001]. Thus far, however, this work is very preliminary: the N-V centers, which are randomly distributed in the diamond instead of being precisely positioned in the cavities, are bystanders in our studies.

Placing Impurities

Many of the experiments on N-V centers to date have been carried out using synthetic diamonds like those used for our optical cavities: the N-V centers were formed naturally but in random locations during the diamond growth process. Now researchers at the Australian National University, the University of Bochum in Germany and Lawrence Berkeley National Laboratory are making great progress in placing individual impurities at specific locations. They use advanced ion implantation techniques to insert single ions of nitrogen with submicron accuracy. Then they heat the diamond to 850 degrees Celsius, which causes the vacancies in the diamond to migrate. When a vacancy meets a nitrogen atom it stays next to it, forming an N-V center.

N-V centers seem a promising technology for processing quantum information, but what about storage for times longer than the millisecond-long decay time of their electronic spin states? Researchers in Mikhail Lukin's group at Harvard University have explored an approach that makes use of the spins of carbon nuclei. Because the nucleus of the most common isotope of carbon, carbon 12, has zero spin in total, the group used carbon 13 atoms, whose nuclei have the spin of their one extra neutron. The scientists transferred the information encoded in a single N-V center spin to a single nuclear spin of carbon 13 and retrieved it 20 milliseconds later. The nuclear spin showed no sign of decay, indicating that the quantum state could survive for seconds. Thus, nuclear spins appear to be a propitious route to qubit storage. The Harvard researchers have also proposed a design for constructing a quantum repeater based on this work. Quantum repeaters are a basic element needed for quantum communication (transmitting qubits over longer distances).

It is an exciting time for quantum information research, with many different computation architectures vying for supremacy. Considering the rapid rise and successes of diamond-based spin research over the past few years and with companies such as Hewlett-Packard getting into the game, the prospect of room-temperature quantum information processors is sounding less like science fiction. The diamond age of quantum electronics could be just around the corner. ■

[THE AUTHORS]

David D. Awschalom, Ryan Epstein and Ronald Hanson are associated with the Center for Spintronics and Quantum Computation at the University of California, Santa Barbara. Awschalom is director of the center and professor of physics and of electrical and computer engineering at Santa Barbara. His research group is primarily concerned with investigating electron-spin dynamics in a variety of semiconductor systems. Epstein obtained his Ph.D. in Awschalom's group, studying nitrogen-vacancy centers in diamond. He is now doing postdoctoral research on trapped ions at the National Institute of Standards and Technology in Boulder, Colo. Hanson was a postdoctoral researcher in the group and has just become assistant professor of physics at the Kavli Nanoscience Institute in Delft, the Netherlands. For his Ph.D., he studied single-electron spins in gallium arsenide quantum dots at the Delft University of Technology.

© 2007 SCIENTIFIC AMERICAN, INC.

WILL THE MEDICINE CABINET
of the future contain experimental drugs?

© 2007 SCIENTIFIC AMERICAN, INC.

EXPERIMENTAL DRUGS ON TRIAL

A controversial lawsuit challenges the FDA's system of controlling access to experimental drugs and, some say, the scientific basis of drug approval

By Beryl Lieff Benderly

Abigail Burroughs was only 21 when she died. If her father and his supporters get their wish, however, she will attain a kind of immortality, joining Brown, Griswold, Roe and Miranda in the band of ordinary citizens whose personal travails have permanently changed the way Americans live.

A lawsuit, *Abigail Alliance for Better Access to Developmental Drugs v. Andrew von Eschenbach*, contends that government regulations kept Burroughs from obtaining potentially lifesaving experimental cancer medicines that her doctor recommended, violating her constitutional right to defend her life. The U.S. Court of Appeals in the District of Columbia ruled against this claim in August, and the plaintiffs plan to appeal to the U.S. Supreme Court [see "Case Study," on next page]. The ultimate outcome could become one of the most important court decisions ever to affect medical science, climaxing a conflict that has simmered since the early days of the AIDS epidemic. On the one hand are the heartrending emotional pleas of dying patients who are eager to try unproved experimental drugs. On the other are scientists determined to preserve the "gold standard" method that for decades has established which medications work—a series of clinical trials that in total can take 10 years to complete. Both sides want to protect patients' lives, but they differ on the best strategies toward that noble end.

The other name on this potential landmark suit is commissioner of the U.S. Food and Drug Administration, the agency that regulates the sale of medicines in the U.S. and the trials that determine which are safe and effective and thus eligible to be marketed. The agency, which is at the same time facing intense pressure to speed drug approvals and also severe criticism for not being careful enough with decisions that could affect public health, declined to comment on the case.

It is clear, however, that an Alliance win could vastly broaden access to medications that have shown only limited evidence of safety and, often, none at all of efficacy. This change would devastate the clinical trial system that has fueled unparalleled medical progress over 45 years, argues the Society for Clinical Trials, one of many opponents. If patients could get experimental drugs outside of rigorous clinical trials, they would be reluctant to bother with such trials, and the surest way to determine whether a drug works and is safe would go out the window.

Hope of being helped motivates many people

KEY CONCEPTS

- Large clinical trials are the gold standard for testing whether experimental drugs work. But they take time and money; meanwhile people who have run out of treatment options may die waiting for results.
- A recent lawsuit challenged the FDA's rigorous requirements for drug testing, with the aim of making experimental drugs much more available.
- But if anyone can gain access at will to experimental drugs—most of which prove useless or even dangerous—will clinical trials be unable to attract subjects? And will scientists end up with no good way to distinguish the truly helpful drugs from the chaff?

—The Editors

STEVEN PUETZER *Iconica/Getty Images*

© 2007 SCIENTIFIC AMERICAN, INC.

CASE STUDY

This fall the U.S. Supreme Court could decide whether to hear arguments about granting dying patients greater access to experimental drugs.

THE CASE

Abigail Alliance for Better Access to Developmental Drugs v. Andrew von Eschenbach (the current FDA commissioner)

THE CONTENTION

FDA regulations kept Abigail Burroughs from obtaining potentially lifesaving experimental cancer medicines that her doctor recommended, violating her constitutional right to defend her life.

THE HIGHLIGHTS

July 2003: Abigail Alliance and the Washington Legal Foundation file suit.

August 2004: U.S. District Court throws out the case, in part because the "court is not persuaded that the plaintiffs seek a recognized [constitutional] right."

May 2006: U.S. Circuit Court of Appeals for the District of Columbia panel rules, 2–1, that a "mentally competent, terminally ill adult patient" has a constitutional right to use, on advice of a doctor, any drug that has passed the first step of the FDA's three-phase approval process [see "Safety in Numbers," on page 97].

March 2007: Full nine-member Circuit Court hears the case reargued.

August 2007: Circuit Court votes, 8–2, against the plaintiffs.

to enter trials, although the studies' entirely different purpose—answering scientific questions about safety and efficacy rather than providing therapy for individual participants—is supposed to be explained before volunteers give their informed consent to participate. So are the statistical unlikelihood of benefit and the real risks of harm from untested chemicals. Research shows, however, that most subjects nonetheless hold what bioethicists call the "therapeutic misconception" that trials aim to cure them and offer a good chance of helping. Even many physicians misinterpret their patients' prospects in trials. Entering a trial really resembles "going to Las Vegas," however, says bioethicist and biochemist Adil Shamoo of the University of Maryland. "Everybody thinks they're going to win in Las Vegas." Most don't.

Long Road to Rx

The quest to protect patients against dangerous drugs is at the root of the modern clinical trial system, which goes back decades. The Food, Drug, and Cosmetic Act of 1938 first required drugs sold in the U.S. to be proved *safe* after 107 people died from the then new wonder drug Elixir Sulfanilamide, which contained the solvent diethylene glycol, a poison used in antifreeze. In 1962 amendments added the requirement that drugs also be proved *effective* after the worldwide thalidomide furor. The experience with thalidomide was a catastrophe in many countries but a triumph for the FDA. Widely used abroad for insomnia and morning sickness, thalidomide caused severe malformations in 10,000 newborns. Because an FDA examiner who was skeptical of the safety data had blocked approval, however, few of those victims were American. Then, 25 years later, after years of pressure from AIDS activists, the FDA began permitting patients with life-threatening illnesses but no approved treatment to take the risk of using experimental drugs outside of trials in limited circumstances. When the alternative is certain death, the FDA acknowledged, the balance of risk and benefit changes. American law has never yet recognized a constitutional right to unapproved drugs, however.

Getting a drug developed and approved takes about eight and a half years, the FDA estimates, and only a small minority of those who start the process finish. The Pharmaceutical Research and Manufacturers of America estimates that five of 5,000 compounds that enter preclinical testing make it to human trials—and only one in five of those may reach the market. The "gold standard" of drug testing, the double-blind controlled clinical trial, compares an experimental drug against the best standard treatment or, sometimes, against an inactive placebo. Neither the subjects nor the health care professionals in contact with them are supposed to know who receives what

ABIGAIL BURROUGHS, who was 21 when she died of cancer in 2001, had sought to join clinical trials of the then experimental drugs Erbitux (cetuximab) and Iressa (gefitinib).

ERBITUX (*above*), marketed by Bristol-Myers Squibb, received approval in 2004; Iressa (*not shown*), made by AstraZenica, was approved in 2003 for treatment of certain cancers.

COURTESY OF FRANK BURROUGHS (*left*); IMCLONE SYSTEMS, INC. (*right*)

© 2007 SCIENTIFIC AMERICAN, INC.

while the experiment is under way. Participants must agree to be randomly assigned to either an experimental arm that receives the as yet unproved medication or to a control arm that does not receive it.

Clinical trials, which occur in three phases, are designed to make statistically valid comparisons of the study arms, not to provide care or to cure subjects. "Clinical research … is not a therapeutic activity devoted to the personal care of patients" but rather an experiment "answering a scientific question, with the aim of producing 'generalizable knowledge,'" write bioethicists Franklin G. Miller of the National Institutes of Health and Howard Brody of the University of Texas Medical Branch in an article in *Hastings Center Report*. Trials "promot[e] the medical good of future patients [through] scientific knowledge derived from experimentation with present patients—a frankly utilitarian purpose."

Phase I introduces the drug into 20 to 80 individuals who are sometimes healthy volunteers. In cases involving highly toxic substances such as cancer drugs, subjects are generally patients with advanced, untreatable disease. This phase seeks to learn about safe dosage levels, side effects, and the drug's effect on metabolism and its mechanism of action in humans. Only occasionally does phase I produce evidence of effectiveness. Drugs that appear sufficiently safe get permission to proceed to the next phase, although that safety finding is not definitive.

Phase II compares the experimental drug and the control using several hundred people who have the condition. It seeks evidence of efficacy, risks and short-term side effects. If the results and the study design appear satisfactory, phase III next tests the drug against a control in up to several thousand people with the condition, seeking much more extensive data about effectiveness and safety. Based on these results, the FDA determines whether the drug merits approval.

Too Long to Wait

Current procedures do permit some patients with serious or fatal diseases but no approved treatment options to obtain, outside of clinical trials, drugs that have passed the second phase. Burroughs's inability to obtain such a promising drug inspired her father, Frank Burroughs, to start the Alliance in November 2001 after her death earlier that year. During her final months, her Johns Hopkins University oncologist thought that Erbitux (cetuximab) or Iressa (gefitinib) could help fight her squamous cell head and neck cancer because both counter epidermal growth factor receptors (EGFR) plentiful in her tumor. Neither had yet won approval.

*** RECALLED ***

ELIXIR SULFANILAMIDE, which contained a poison found in antifreeze, killed 107 people in 1937. The tragedy spurred the passage of the 1938 Food, Drug, and Cosmetic Act, which required that manufacturers prove drug safety to the FDA before receiving marketing approval.

To be permitted to experiment on human beings in the first place, sponsors of candidate drugs must submit data from preclinical studies showing a reasonable likelihood they will prove safe and effective. The sponsor must also recruit the experimental subjects, who, in the case of serious or fatal conditions, are often individuals such as Burroughs, with advanced disease and no approved treatment options. A clinical trial offers them a shot at getting an experimental drug, but how good a shot depends on the phase of the trial they participate in and the study's research design. The sponsor provides the test drug free to participants but is under no obligation to accommodate others out of its often small supply.

That, Scott Gottlieb believes, is why an Alliance win may increase access less than many people expect. Gottlieb is a former FDA deputy commissioner for medical and scientific affairs and now a physician in private practice and a resident fellow at the American Enterprise Insti-

Most subjects hold what bioethicists call the "therapeutic misconception" that trials aim to cure them and offer a good chance of helping.

COURTESY OF FDA HISTORY OFFICE

© 2007 SCIENTIFIC AMERICAN, INC.

*** RECONSIDERED ***

THALOMID (THALIDOMIDE), manufactured by Celgene, received approval in 1998 to treat an inflammatory condition of Hansen's disease, or leprosy. In the early 1960s a skeptical FDA examiner declined to approve thalidomide, sparing Americans from the number of drug-related birth defects experienced in countries where the drug was used to combat morning sickness.

*** DENIED ***

LAETRILE, sold in other countries as a cancer treatment, is not approved by the FDA, because it lacks proof of efficacy and produces side effects resembling cyanide poisoning.

tute. "The biggest impediment [to early access outside of trials] is the unwillingness of some companies to offer the drug," he says. A group of patients with Parkinson's disease, for example, lost a suit to force Amgen Corporation to continue providing them with an experimental drug they believed to be very effective that they received as subjects in a trial the company chose to terminate. Sponsors may stop trials at any time. The FDA also stops trials when the pattern of "adverse events" indicates undue risk or when the evidence of benefit is so strong that maintaining a control group becomes unethical.

The Iressa trial had no room for Burroughs, and the Erbitux trial was studying only colon cancer. Neither sponsoring company would supply her under the rules permitting compassionate use of experimental drugs outside of trials. She died soon after finally securing a spot in a trial of a third EGFR inhibitor. Both drugs she sought subsequently gained approval; Erbitux is now widely used for her cancer. Abigail Alliance's cofounder and senior adviser, professional geologist Steven Walker, has seen the benefits of using such drugs. Walker's wife took Erbitux in a trial in September 2002 and was "resurrected," changing from "a bedridden, dying person into someone who was skiing, hiking and working every day," he says. After "marginal progression [of the disease] in her liver," however, she was no longer given the drug and died in June 2003.

A Better System?

The dispute ultimately involves a "fundamental trade-off" over "how … patients are best served in the long run: through earlier access to promising drugs that could compromise the ability to conduct very formal and rigorous clinical studies or through more rigorous evaluations that might forestall early access but enable the development of rigorous clinical data that can guide future decision making," states Gottlieb in an essay in a February 2007 white paper published by the Food and Drug Law Institute, a nonprofit educational organization in Washington, D.C.

Instead of sound science, Shamoo believes, open access to experimental drugs would create "large, uncontrolled clinical trials" that would endanger patients by exposing them to untested chemicals and would slow research by removing subjects' incentives to enter the controlled clinical trials capable of producing valid results. What is more, people would pay for the drugs they received outside of trials, giving sponsors incentives to "sell the unapproved drug today" rather than financing the "scientifically rigorous and expensive clinical trial process," the FDA argued in its brief for the Appeals Court en banc hearing. Letting everyone choose any drug could even harm promising medicines, critics believe, because very sick people using inappropriate chemicals might suffer adverse reactions serious enough to convince the FDA to reject drugs that could otherwise help carefully chosen patients.

Walker, however, says that open access will "end the abandonment" of the great majority of patients who cannot get into trials. He foresees no damage to recruitment because, under the Alliance's proposal, anyone seeking access to an experimental drug would first have to apply for a clinical trial. Nor would patients opt to try dangerous or ineffective chemicals, Walker in-

COURTESY OF FDA HISTORY OFFICE

© 2007 SCIENTIFIC AMERICAN, INC.

sists, because "patients and physicians do not pursue investigational drugs in any kind of significant number unless that drug has in fact shown some evidence of effectiveness and at least enough safety … that the patient's physician" recommends trying it.

Perry Cohen, founder and director of the Parkinson's Pipeline Project, a patient group working to "accelerate innovation" in treatment, also asserts that patients can make good choices. Cohen, who was forced by the disease to retire as a health consultant, now serves as an FDA patient representative for Parkinson's. "Even if you're not going to die," he says, people with "a serious illness leading to a worse fate" deserve the right "to make the risk-benefit trade-off individually, with proper counseling from expert medical professionals" and to decide whether to try experimental drugs.

Critics reject those contentions. Many side effects are unknown before large-scale testing, they argue. Once a right is established for some, they add, it inexorably expands to others.

Allowing people to choose among unapproved chemicals would not mean that "the patient would suddenly have access to the one drug out there" right for him and his condition, says Colin Begg, chair of epidemiology and bio-

*** APPROVED ***

PENICILLIN, the "wonder drug," was tested by the FDA for wartime use and later certified for public consumption.

*** CAVEAT EMPTOR ***

Cooper's New Discovery and Mrs. Winslow's Soothing Syrup, pre-1906 "patent" medicines that predate the FDA, were typical in their extravagant claims.

*** APPROVED *** **HUMULIN (HUMAN INSULIN), marketed by Eli Lilly, became the first genetically engineered drug approved by the FDA in 1982.**

SAFETY IN NUMBERS

New drugs must pass three clinical trial milestones before they can be marketed to the public. A small fraction of candidates complete the development and approval process, which takes about eight and a half years.

Phase I: Tests for safe dosage levels and side effects in 20 to 80 volunteers. Also probes the drug's effect on metabolism and its mechanism of action.

Phase II: Several hundred subjects receive the drug in a probe of its efficacy, risks and short-term side effects.

Phase III: More extensive data about effectiveness and safety come from testing the drug in up to several thousand people.

COURTESY OF FDA HISTORY OFFICE

© 2007 SCIENTIFIC AMERICAN, INC.

statistics at Memorial Sloan-Kettering Cancer Center in New York City and of the Society for Clinical Trials committee that drafted a position paper on the Abigail Alliance proposals. The patient "would have access to an entire cupboard full of drugs, most of [them] useless or toxic, [without] any idea of which … to take," because physicians not expert in the drug would have no evidence to guide the decision.

The trade-off, however, need not be as stark as the most extreme positions imply, Gottlieb believes. Improvements in trial design could get answers sooner, shortening everyone's wait for new drugs and reducing the number of participants needed for control arms, he says. He sees "a lot of opportunity to develop better scientific tools for measuring drug response and [for designing] clinical trials using more adaptive approaches." The "gold standard is crumbling," Cohen adds. "Most of the acute conditions that fit that model are [already] done." Like Walker, Cohen wants more emphasis on earlier approval followed by extensive postmarket monitoring. Modern information technology, adaptive research designs and Bayesian statistics can make this feasible, he believes. (Ironically, the FDA is meanwhile facing criticism for giving overly speedy approval to drugs such as Vioxx, for pain, and Rezulin, for diabetes, which have been withdrawn from the market as unsafe.)

Bayesian statistics do not require "perfect experiments," Gottlieb explains, but rather use "real-world data, 'dirty data' … to derive conclusions about probabilities from nonrandomized data sets." Adaptive trial design permits changes to be made while a trial is under way based on the data that emerge during the trial. Researchers, for example, could modify randomization to give more subjects a treatment apparently producing a better response or to overenroll categories of patients who, because of their individual characteristics or responses, appear likeliest to benefit from the experimental drug. Gottlieb, Walker and Cohen also all favor increased use of surrogate end points, which substitute physiological markers believed to predict outcome for evidence of actual outcomes. Such approaches can "speed drug development and … make clin-

Proponents of greater availability for new drugs want more emphasis on earlier approvals followed by extensive postmarket monitoring.

SHORTENING THE CRITICAL PATH

Although clinical trials are still the FDA's gold standard for assessing the safety and effectiveness of drugs and medical devices, the agency acknowledges plenty of room for improvement in testing methods for prospective therapies. Its Critical Path Initiative, launched in 2004, is an attempt to modernize the process at every stage on the path from promising compound to approved product, in the hope of improving data quality and speeding the best drugs to market. A list of 76 priority projects emphasizes applying the latest scientific tools and techniques to ruling out useless or toxic compounds early and assessing more quickly whether candidate drugs have the desired effect on disease.

Causing liver injury, for instance, leads many drugs to fail in early human trials, and even approved medicines must sometimes be pulled from the market for the same reason. To avoid these late and costly realizations, the FDA is collaborating with a computer-modeling firm to design a "virtual liver" that can help predict whether a compound may be toxic in people, alone or in combination with other drugs. Many Critical Path projects are also exploring the use of biomarkers, which are various forms of evidence of biological activity inside the body—such as levels of certain proteins, gene activity or internal imaging data—to gauge whether a drug is having the predicted effect without having to await results of a large human trial. This year, for example, the FDA permitted the first "phase 0" human trials to look for biomarker evidence of a candidate drug's action on its target in the body. In the earliest of these, National Cancer Institute researchers gave tiny doses of a compound called ABT-888 to six advanced cancer patients and saw that their levels of a significant protein dropped by more than 90 percent—a sign that the FDA and the drug's developer can use to accelerate the compound's progress to formal clinical testing.

Making those trials shorter and smarter is another agency goal, and approaches under consideration include adaptive design, which allows protocol changes in response to preliminary data while a trial is still under way. Testing drugs in multiple combinations, as they are likely to be used in reality, could also speed approvals. Plugging data from early trials into computer simulations can help identify the most effective drug doses for later, larger trials. The FDA is also encouraging pharmaceutical companies to pool their data from every stage of drug testing and use bioinformatic techniques to extract the maximum amount of information from that raw material, so that problems with a whole class of compounds are flagged sooner, as are promising leads to valuable new medicines. ***—Christine Soares, staff editor***

[THE AUTHOR]

Beryl Lieff Benderly is a prize-winning health and science journalist in Washington, D.C. Her books include *In Her Own Right: The Institute of Medicine's Guide to Women's Health Issues* (National Academies Press, 1997).

COURTESY OF FDA HISTORY OFFICE

© 2007 SCIENTIFIC AMERICAN, INC.

JOSH SHER (*Vioxx box*) AND JOHN KAPRIELIAN (*Vioxx pack*) *Photo Researchers, Inc.* (*left*); JUSTIN SULLIVAN *Getty Images* (*right*)

ical trials more targeted to patients who are likelier to benefit," yielding "much more reliable information about who's going to ultimately benefit from using the drug," Gottlieb says. The FDA is currently evaluating these techniques [*see box on opposite page*].

But "there's no promised land" in that direction, Begg warns. He believes that Bayesian methods are relatively minor adaptations of the way patients are randomized and the way trials are designed. "They do not get the answer faster; they result in a longer trial," he says. Surrogate markers do not always reliably predict clinical outcome. Nor will these techniques resolve the assertions of the Abigail Alliance, Begg continues: "It's just a different way of designing clinical trials. You still have to do the efficacy testing [on] patients to figure out which drugs work and which drugs don't."

"The long history of medical research is replete with treatments that initially seemed promising to patients, doctors, and especially to their own inventors, but which careful study revealed to be worthless or harmful," states the Society for Clinical Trials position paper. In the cancer field alone, these failures include laetrile, initially impressive in animal studies, and high-dose chemotherapy followed by autologous bone marrow or stem cell transplant for breast cancer. Thousands of women underwent, and some died from, this excruciating and costly experimental procedure after a lawsuit forced insurers to pay but before clinical trials finally proved it no more effective than standard therapy.

Other Implications

The courts will decide not on issues of experimental design and risk-benefit analysis but on issues of legal precedent and interpretation. An Alliance win, suggests John A. Robertson of the University of Texas at Austin School of Law, could, for example, overturn rules or laws against creating embryonic stem cells. "Under the principle recognized in *Abigail Alliance*," he writes in an article in the *Hastings Center Report*, courts would have to weigh "whether the state could protect embryos at the expense of the patient's life and liberty."

Even if the FDA prevails, some change appears likely. The agency is working on new rules it says will make early access outside of trials less onerous. "Scientists have to realize that there's a crisis of confidence" in clinical trials among the public, Cohen says, and "holding up this standard as immutable is part of" the reason. Whatever the *Abigail Alliance* case brings, the health of the clinical trial system, and thus of the American people, depends on restoring that confidence. ■

SCI AM

More on the Web

To see more historic drugs, go to **www.SciAm.com/ontheweb**

* WITHDRAWN *

VIOXX (ROFECOXIB), an anti-inflammatory approved in 1999 that soon became widely prescribed, was in 2004 withdrawn from the market by developer Merck & Company in the wake of concerns about the risk of heart attack and stroke in patients.

* ALERT *

AVANDIA (ROSIGLITAZONE), made by GlaxoSmithKline, was approved in 1999 to treat type 2 diabetes; in May the FDA issued a safety alert about its potential to raise the risk of heart attack, but it has not yet confirmed the clinical significance of related findings. A drug for type 2 diabetes in the same class, Rezulin (troglitazone), was withdrawn voluntarily in 2000 by the Parke-Davis division of Warner-Lambert at the FDA's request because of concerns about its risk of liver toxicity.

MORE TO EXPLORE

Access before Approval—A Right to Take Experimental Drugs? Susan Okie in *New England Journal of Medicine,* Vol. 355, No. 5, pages 437–440; August 3, 2006.

The Right to a Trial. Jerome Groopman in the *New Yorker;* December 18, 2006. **www.newyorker.com/fact/content/articles/061218fa_fact**

CDER Handbook, a description of the drug development and approval process provided by the Center for Drug Evaluation and Research of the Food and Drug Administration: **www.fda.gov/cder/handbook/index.htm**

FDA press release on the proposed new regulations: **www.fda.gov/bbs/topics/NEWS/2006/NEW01520.html**

© 2007 SCIENTIFIC AMERICAN, INC.

Big Lab on a Tiny Chip

Squeezing a chemistry lab down to fingernail size could provide instant medical tests at home and on the battlefield

By Charles Q. Choi

KEY CONCEPTS

- Researchers are devising tiny, portable chips that could rapidly detect pathogens or biological weapons in a sample of a person's blood.
- Microfluidics—using air pressure or electricity to move droplets through tiny chemical reactions—is the key to making labs-on-chips practical.
- The University of Michigan at Ann Arbor has devised a chip that could test for influenza, but it still requires an external air supply. The Massachusetts Institute of Technology has crafted an electric conveyor that could efficiently drive microdroplets.
- Ultimately, consumers could have labs-on-chips in their home for fast diagnosis of common illnesses. —*The Editors*

Imagine shrinking the beakers, eyedroppers, chemicals and heaters of a chemistry lab onto a little microchip that could dangle from a key chain. A growing number of companies and universities are claiming to have devised such marvels, ready to perform vital analyses from detecting biological warfare agents in a soldier's bloodstream to identifying toxins in a tainted package of hamburger meat. Almost all the new devices are surprisingly far from portable, however. The sensor that examines a drop of blood or speck of beef might indeed fit in one's hand, but the equipment required to actually move a fluidized sample through the chip's tiny tubes often occupies a desktop or more.

Two research teams are overcoming that hurdle with creative microfluidics—the precise manipulation of microscopic droplets. By moving liquid molecules with air or electricity, the groups are integrating the equipment needed to sample, analyze and report, all on a fob the size of a USB flash drive. And although the current chips are being crafted by hand, the designs could ultimately be mass-produced. That prospect would finally bring labs-on-chips to the places they are most desirable—the developing world, the battlefield and the home—where they could quickly detect HIV, anthrax or *Escherichia coli*. A chip could even be implanted into a diabetic's body to help monitor the person's glucose and insulin levels.

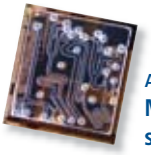
Pushed by Air

As a tool, labs-on-chips have become increasingly popular among researchers because they can conduct hundreds of experiments simultaneously at a mere fraction of the time, space and cost of long-standing benchtop processors. Tiny channels and valves inside the chips can heat, cool or mix small samples and reagents, as well as enable more exotic tests such as electrical stimulation. The surrounding apparatus required to perform the internal heating, cooling and mixing is often comparatively bulky, however, because of the unique behavior of fluids. When trapped inside incredibly narrow tubes, even watery compounds behave like syrup: they are difficult to push around. And ironically, when they do flow, they are remarkably free of turbulence, making it hard to mix them with reagents for chemical reactions. Forcing liquids through the chips with compressed air requires bulky plumbing. Electrically driving the liquids requires high-voltage power supplies.

The tests these desktop labs-on-chips can handle have gotten progressively more sophisticated. In 1998 chemical engineer Mark Burns and geneticist David Burke of the University of Michigan at Ann Arbor demonstrated the first chip that could identify a particular gene or variation of it. Since then, the researchers have

GEORGE RETSECK

© 2007 SCIENTIFIC AMERICAN, INC.



ACTUAL SIZE: **The University of Michigan's influenza chip measures 1.5 by 1.6 centimeters.**

steadily miniaturized and integrated the surrounding equipment. "The vision," Burns says, "is to get genetic analysis equipment to whoever wants it, including people at home." For example, a concerned parent could quickly determine if a child who becomes ill at 2 A.M. might have influenza, instead of hauling her to an emergency room and having to wait for results from a lab.

Analyzing the genes in a droplet begins with amplification: heat and the addition of particular enzymes facilitate the creation of millions of copies of the genetic material. Those clones are then mixed with digestive enzymes that locate and snip out specific DNA sequences. Fluorescent dye molecules attach to the DNA snippets. Electric fields then move the snippets through a gel, a process known as electrophoresis. The speed at which the snippets move, which depends on their size and electric charge and is observed with light, reveals details about the DNA, such as whether it matches segments from deadly germs or harmless ones.

Burns's work had actually begun in 1993, and by 1998 his group had devised ways to propel droplets down channels using small puffs of air, "much like one would blow liquid through a straw," Burns explains. Further design enabled temperature sensing, amplification, electrophoresis and fluorescence detection to all take place on a wafer roughly half a centimeter wide and three centimeters long—about the size of a half-stick of chewing gum.

The equipment that drove those operations still surrounded the wafer, however. Burns and Burke spent the next seven years miniaturizing it. Along the way, the duo replaced the original DNA amplification method (strand displacement amplification, or SDA) with the increasingly popular technique known as polymerase chain reaction (PCR). PCR required only one enzyme instead of two, greatly simplifying the onboard chemistry. But PCR also required extraordinarily intricate changes in temperature. For SDA, the chip needed to heat the sample to 50 degrees Celsius and keep it there. PCR demanded a cycle of heating and cooling, from 90 degrees C down to 50 and back up to 70, which had to be repeated 35 times. The researchers had to hit more than 100 unique temperature points, which had never been attempted with anything nearly as portable as the device Burns and Burke envisioned.

Furthermore, the spot where each step took place had to be thermally isolated, a huge feat when all the processes are crammed onto a microchip. Burns and Burke spent months experimenting with different materials, along with gate patterns and valves—shaping channel junctions and testing coatings so they could move and mix molecules and siphon away excess fluids caused by certain reactions.

By 2005 the team had integrated the electronics, sensing, heating and electrophoresis components onto a silicon wafer about the size of a quarter. A glass substrate housed the liquid channels. Instead of requiring many pressurized air connections to supply the pneumatic force needed to open and close each valve and push fluid around, the chip required just two conduits, significantly reducing its bulk. A single air supply became feasible after Burns and Burke found they could make the many valves with wax. Electronics on the chip would heat each valve individually at the right time during analysis, making the wax mobile enough for air pressure to push the values open or closed.

The team named their genetic analysis chip "VIPER" (*v*alved, *i*ntegrated *P*CR *e*lectrophoresis *r*estriction digest), and its function is to distinguish variants of influenza virus genes [*see illustration on next page*]. The chip can perform genetic analysis in as little as 15 minutes, 10 times faster than PCR takes in a standard lab. As Burns points out, "Other diseases could be detected by merely changing the liquid reagents, or 'wetware,' just like many programs can be run on a computer by changing the software.... The size suggests that mass production could be relatively cheap, under $1 per chip."

The primary remaining hurdle is the source of air pressure, which is still external to the chip. The system could use a low-pressure carbon dioxide cartridge about the size of a finger, as well as off-chip electronics, to distribute the air power to various valves. The other off-chip compo-

MARK BURNS University of Michigan (chip)

CHIPS IN DEVELOPMENT

HIV TEST
Harvard-M.I.T. Health Sciences
Disposable wafer the size of a business card would test for HIV and read out immediately; intended for poor countries.

MRSA TEST
Cepheid
Chip for hospitals would sense methicillin-resistant staph bacteria to reduce the spread of hospital-acquired infections.

CANCER SCREEN
University of Alberta
Wafer would perform "fluorescent in situ hybridization"—a screening test that can detect the chromosome mutations of various cancers.

© 2007 SCIENTIFIC AMERICAN, INC.

The tiny size "suggests that mass production could be relatively cheap, under a dollar per chip."
—Mark Burns

nent is a blue LED that illuminates the genetic material during electrophoresis. Burns says these pieces could be shrunk onto a second chip, with both chips fitting onto the equivalent of a USB flash drive. "Getting to the USB drive size gives me claustrophobia," Burns says, "but it is possible. It would be easier to make an iPod-size device, giving room to put a small compressed air cylinder inside, as well as multiple chips."

Private money may be needed to realize either configuration; Burns and Burke do not have the funding to develop them. Yet biomedical engineer David Beebe of the University of Wisconsin–Madison thinks that Burns and Burke "have made the most convincing argument that a portable lab-on-a-chip might actually be possible in a commercial way." Perhaps the investment will come.

Propelled by Electricity

As Burns's story illustrates, the greatest challenge in devising labs-on-chips may be moving the liquids onboard with as little power as possible. By tweaking the design of a tiny pump, researchers affiliated with the Massachusetts Institute of Technology's Institute for Soldier Nanotechnologies have taken a major step toward solving that problem.

Martin Bazant, an applied mathematician, leads the M.I.T. work, and he has chosen electricity as his propellant, not pressurized air. "There are no moving parts, making things simpler," he says. The pump his team is developing requires only a few volts, which a watch battery could provide.

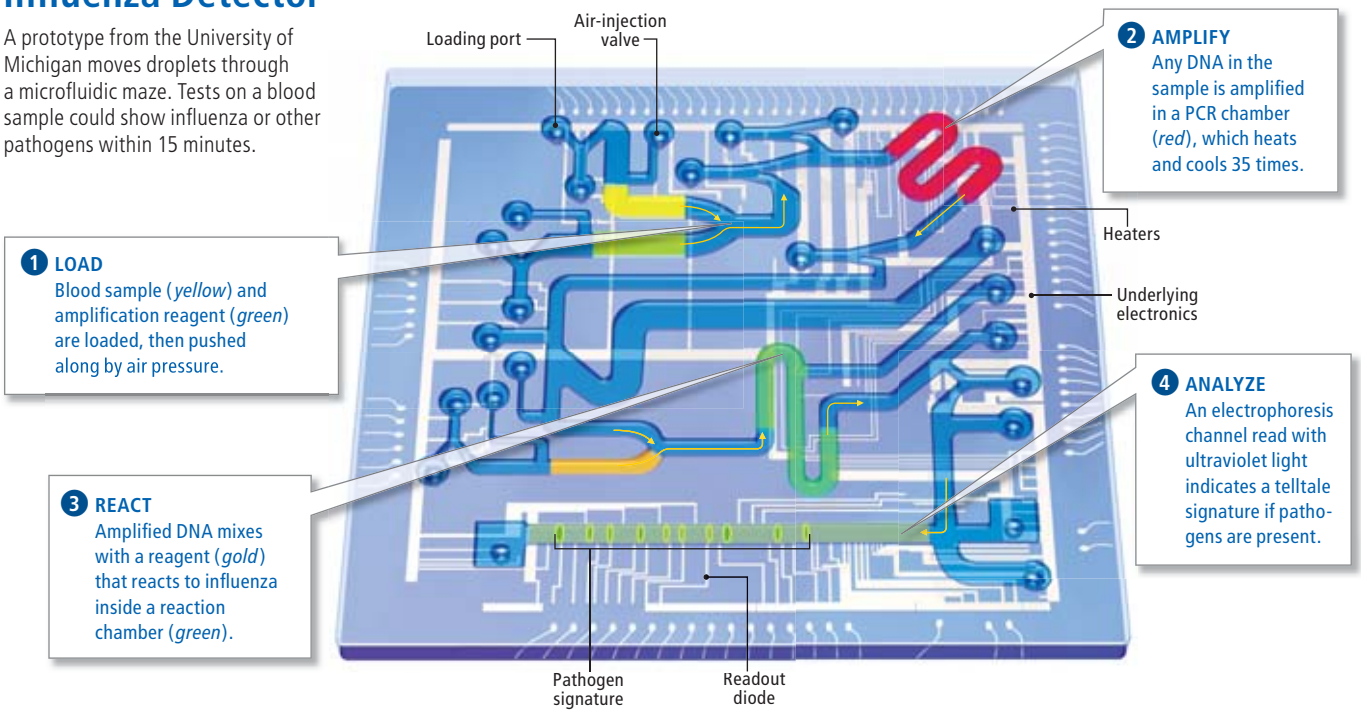
Bazant studies electroosmotic flows, wherein electric fields drive charged molecules in solution. For years electroosmotic pumps operated on direct current (DC) of 100 volts or more, untenable for a portable lab-on-a-chip. In 1999 scientists began reporting alternating current (AC) electroosmosis; AC systems require far lower voltages because they can use multiple electrodes spaced throughout a channel, whereas DC systems rely on one big electrode placed at each end of a channel.

An AC design can be visualized as a length of railroad track. The low, flat railroad ties are electrodes, and fluid covers them between the rails. If electrodes alternate between positive and negative charges, liquids flow in one direction. At first glance this seems impossible; a fluid molecule should simply oscillate back and forth between two given electrodes, resulting in no net movement. But introducing irregularities in electrode shape, spacing or coatings causes fluids to prefer one direction.

The problem for labs-on-chips, however, was that the AC pumps moved fluids too slowly. In 2003 Bazant and his colleagues theorized that electrodes shaped like a set of two stairsteps (if viewed on end) would lead to complex three-di-

Influenza Detector

A prototype from the University of Michigan moves droplets through a microfluidic maze. Tests on a blood sample could show influenza or other pathogens within 15 minutes.



GEORGE RETSECK; SOURCE: UNIVERSITY OF MICHIGAN AT ANN ARBOR

© 2007 SCIENTIFIC AMERICAN, INC.

mensional flows that might generate quicker net movement. This, too, seemed counterintuitive; wouldn't a stairstep interfere with flow?

Perhaps so. But Bazant realized he could use the interference to his advantage. AC designs were slow because the alternating current made fluid molecules slip backward between the flat electrodes and oppose the overall progress. The stairstep, however, should cause back-slipping molecules between adjacent electrodes to circulate in a loop, creating a local eddy. A series of eddies between a string of electrodes could behave like the rollers under a conveyer belt, helping the fluid above them move along [*see box at right*].

In 2006 Bazant's team presented an AC electroosmotic conveyor that flowed more than 10 times faster than prior designs at the same voltage. The flow rates "approach the flows seen in [air] pressure-driven systems," Bazant says. Furthermore, when his group tested a prototype conveyor, the researchers found that molecules in the eddies stayed trapped there only briefly, with most diffusing out in milliseconds. In a chip, that circulation would ensure that all the molecules in a sample would advance downstream and undergo reactions, dispelling concerns that certain target molecules could be missed if they stayed stuck in the eddies.

One limitation is that ions could crowd one another near an electrode surface, hindering flow. But Bazant thinks he could overcome this problem by diluting samples, adding molecules that break up crowding, or giving the electrodes special water-repellant coatings. His team is just beginning the laborious task of adding microfluidic-analysis components to its conveyor for a complete lab-on-a-chip. "Our goal," he says, "is to deliver to the [U.S.] Army a device the size of a wristwatch that can look for specific signals in the saliva or blood, such as messenger RNA sequences, that mark the body's response to exposure to a wide variety of biowarfare agents."

Worth the Bother?

Although truly portable labs-on-chips seem inevitable, some critics wonder if anyone will want them. In the U.S., physicians deliver patient samples to off-site testing labs that are cost-effective and efficient. Hospitals have labs on-site. "And for most tests, you don't need answers right away," Beebe admits.

Nevertheless, large markets loom. Labs-on-chips could fulfill many uses at home. They could also improve medical diagnostics throughout the developing world, where conventional labs are rare and where doctors often see patients once and then never again, "so you better give them an answer while they're there," Bazant says.

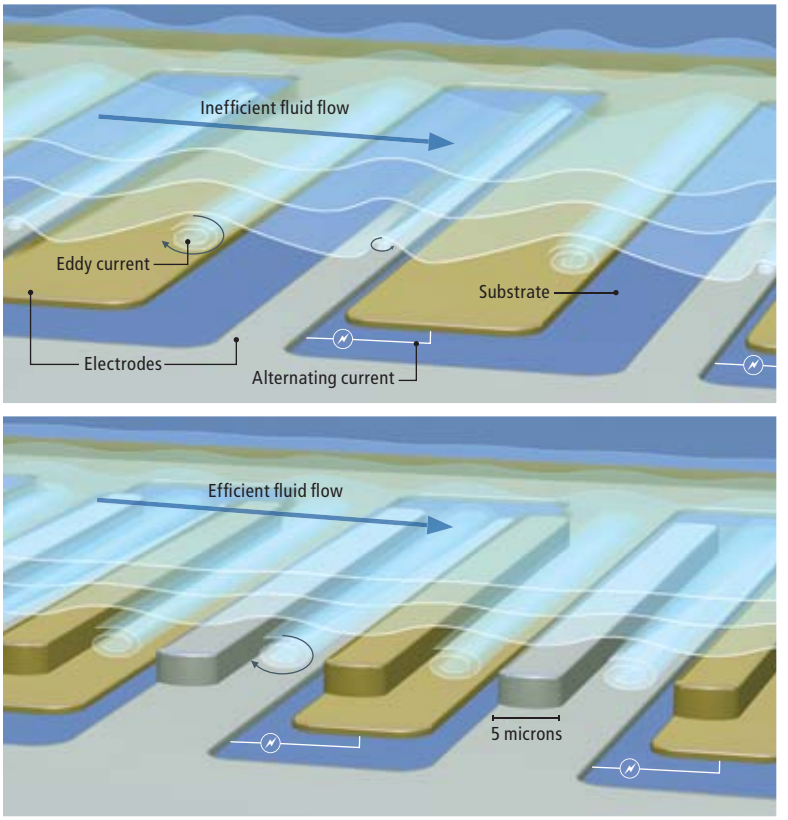
Medics on battlefields cannot just send a sample to a lab either. And even in suburbia, Bazant notes, "you can imagine paramedics and other first responders getting answers on patients right away," which could improve survival rates at the scene or during treatment in an ambulance. If made small enough, labs-on-chips might even be implantable; attached to a tumor, a chip could track whether it is growing or how drugs are affecting it. So the designers push ahead. "We want to show what's possible," Burns says. "We want to show how intelligent and powerful these devices can become." ■

Charles Q. Choi, a frequent contributor, wrote about the International Polar Year in the March issue; he has traveled to seven continents.

[NEW DESIGN]

Microconveyor

Alternating current along a string of electrodes can pump liquids along a microfluidic channel. But turbulence between electrodes (*top*) makes the net progress slow. A novel pump design from M.I.T. (*bottom*) speeds flow by a factor of 10; shaping each electrode like a step creates eddies that act like rollers in a fluid conveyor belt.



GEORGE RETSECK

MORE TO EXPLORE

An Integrated Microfluidic Device for Influenza and Other Genetic Analyses. R. Pal et al. in *Lab on a Chip,* Vol. 5, No. 10, pages 1024–1032; October 2005.

Fast AC Electro-osmotic Micropumps with Nonplanar Electrodes. John Paul Urbanski, Todd Thorsen, Jeremy A. Levitan and Martin Z. Bazant in *Applied Physics Letters,* Vol. 89, pages 143508-1–143508-3; October 2, 2006.

© 2007 SCIENTIFIC AMERICAN, INC.

The Trouble with Men

Deadbeat granddads, life-shortening sons and genetically bullying brothers—these are just a few effects revealed in biologist Virpi Lummaa's studies of how evolutionary forces shape later generations BY DAVID BIELLO

Sons are tough on their mothers. Whether it is heavier birth weights, amplified testosterone levels or simple, hair-raising high jinks, boys seem to take an extra toll on the women who gave birth to them. And by poring over Finnish church records from two centuries ago, Virpi Lummaa of the University of Sheffield in England can prove it: sons reduce a mother's life span by an average of 34 weeks.

The 33-year-old Finnish evolutionary biologist, aided by genealogists, has scoured centuries-old tomes (and decades-old microfiche) for birth, marriage and death records—and clues about the influence of evolution on human reproduction. Historians, economists and even sociologists have long used such tactics to explore their fields, but Lummaa is among the first biologists to enlist *Homo sapiens* as an animal whose population can be followed over time.

After all, humans are relatively easy to track and offer the signal advantage of occasionally keeping detailed records. "I always wanted to work on primates," Lummaa says. "But if I wanted to collect a similar data set on wild chimps, I would be struggling. I've decided to study another primate in the end." Of late, one of her subjects has been premodern mothers among the Sami people of Finland, who are famous for their reindeer herding.

Among this group, she found that those who bore sons had shorter life spans than those who gave birth to daughters. This discrepancy has to do with birth weight—male babies are typically larger—as well as testosterone. "Testosterone can compromise your immune system; it can affect your health," Lummaa says, and the mothers of sons proved especially susceptible to endemic infectious disease, such as tuberculosis. "Boys are a little bit more costly" to raise than girls as well, because they drain more physical resources from their mothers, she adds, as has been seen in other mammals, such as the red deer. Sons also are not as likely as daughters to stick around to help their mothers out later in life.

More recently, Lummaa and her colleagues have been studying how sons are not just tough on their mothers but also hard on their siblings. Those born after a son were physically slighter, had smaller families and generally had a greater chance of dying from an infectious disease. The effects held up whether the elder

VIRPI LUMMAA

HISTORICAL BIOLOGY: Pores over records of preindustrial Finns to study reproductive behavior from an evolutionary perspective.

BOYS ARE BAD: Finds that sons shorten mothers' lives and decrease reproductive fitness of younger siblings; grandfathers do not help children survive, either.

ON HER RESEARCH APPROACH: "You need to have long-term data, and you need to have it over several generations. There isn't such data from many animal species."

ESKO PETTAY

© 2007 SCIENTIFIC AMERICAN, INC.

Send us your answer and you could win $1,000.

What the Rabbit Corkscrew said to the wine bottle is strictly up to you. Your answer may or not be based on features of the Rabbit Corkscrew (below). Make it witty, imaginative, surprising, clever or profound. And a little "Rabbit-y."

To Enter: Send your answer with return address to Rabbit Contest, Metrokane, Inc., 150 East 58th Street, New York, NY 10155. Or enter at metrokane.com. No purchase necessary. Only one entry per household. The 5 best answers win $1,000; 25 runner-ups win a Rabbit Corkscrew. Contest ends January 1, 2008.

Rabbit Features: Pulls a cork in 3 seconds flat; Tested for 20,000 cork pulls*; Works on all bottle sizes and plastic corks; All-metal gears; 10-Year Warranty.

* Assumes replacement of spiral after 800-1,000 cork pulls

Where To Go Rabbit Hunting:
Bed Bath & Beyond, Beverages & More, Carson Pirie Scott, Corkscrew.com, Crate & Barrel, IWA.com, Fortunoff, Macy's, Sur La Table, Total Wine & More.

the original rabbit
metrokane

brother died in childhood or not, suggesting that the negative outcome is not a result of some direct sibling interaction, such as competition for food, regular beatings or the practice of primogeniture, in which the eldest brother inherits everything. "Big brothers are bad for you," Lummaa explains. "If the fifth-born was a male, then the sixth-born is doing worse."

This phenomenon is particularly evident in twins where one is male and the other is female. Of 754 twins born between 1734 and 1888 in five towns in rural Finland, girls from mixed-gender

OLD BIRTH AND DEATH RECORDS from a Lutheran parish in Finland give evolutionary perspective to reproductive behavior.

pairs proved 25 percent less likely to have children, had at least two fewer children, and were about 15 percent less likely to marry than those born with a sister. This brotherly influence remained the same regardless of social class or other cultural factors and even endured if the male twin died within three months of birth, leaving the female twin to be reared as an only child.

The reason that the female half suffers, Lummaa speculates, is because of testosterone exposure in the womb. Researchers have seen such hormonal influence in other animals, including lab rats and cows. When a cow has mixed-sex twins, the female is occasionally born sterile because of testosterone influence.

Whatever the cause, there is no question of the outcome: mothers of opposite-sex twins end up with 19 percent fewer grandchildren than moms of same-sex twins, meaning evolution would seem to favor the latter. "Biological differences between male and female are not determined [simply] by the chromosomes one inherits at birth," says anthropologist Christopher Kuzawa of Northwestern University. The sibling effects "impact reproductive success and thus have evolutionary significance."

The results are somewhat puzzling, comments Kenneth Weiss, a biological anthropologist and geneticist at Pennsylvania State University. He notes that "if twinning is genetic, then there should be a slight selection bias against it, so that twinning would be kept rare. But some animals twin routinely." Given the seeming conflict, he says, "there are dangers in overinterpreting 'fitness' effects, even if the observation is correct."

That danger is especially acute when trying to apply the results to the present day. Access to effective birth control, an abundance of food, and low child mortality rates would all obscure the evolutionary influences seen in the preindustrial data. "It's almost a shock when you realize that 100 to 150 years ago, 40 percent of babies died before they reached adulthood," even when adulthood was defined as age 15, Lummaa notes.

But most of the world's population still lives under similar conditions to those faced by preindustrial Finns. "Who gets the most kids and who puts forward the most genes are still going to be the people who make up the next generation," Lummaa says. "There is no reason why the principles of evolution wouldn't apply." And she hopes to test her Finnish findings against more modern demographic information, such as the ongoing collection of health records for families in the Gambia, on Africa's west coast.

Lummaa and her colleagues have also begun to parse the Finnish records to understand grandparents and the evolutionary conundrums they pose. Her group's previous research has shown that grandmothers provide direct aid in ensuring the survival and reproduction of their grandchildren. The same records revealed,

ESKO PETTAY

© 2007 SCIENTIFIC AMERICAN, INC.

however, no such benefit from fathers and grandfathers. Whereas having a father around did seem to aid children in getting married earlier, a living papa did not increase the number of grandchildren. "If anything there's a negative effect," she concludes. This could be because of the cultural tradition of catering to men, particularly old men. "Maybe if you had an old grandpa, he was eating your food," she speculates.

Or, possibly, longevity in men is simply a by-product of selection for longevity in women. And it could be that because men can reproduce throughout their lives, they are less vested in anyone other than their own children; Lummaa is examining whether men continue to procreate into old age, although a monogamous culture such as Finland's argues against it. "Men past 50 had a chance of finding someone who wanted them," unlike a woman past childbearing years, Lummaa says. "What, if any, benefits do men get from reaching old age?"

Lummaa and her colleagues are also using the data to explore issues of class, showing that the rich in olden times produced more heirs than the poor. Populations in the richest parts of the world seem to have reversed that long-standing trend. "Perhaps under the current circumstances, we are investing in quality over quantity," Lummaa speculates. "The satisfactory answers are still kind of missing."

But the deleterious effects of males—and the benefits of grandmothers—stand out clearly. This does not bode well for Lummaa, who gave birth to a son, Eelis, in March. "I can certainly see that it's taking a lot of energy, and I'm sure it's aging me," she chuckles. "How on earth these women managed to give birth every year is truly amazing."

And certainly there will be no shortage of sons, despite their costs. "If you produce a really, really good son, he can produce a lot of offspring," Lummaa notes—the best outcome from an evolutionary standpoint. "You might lose more by producing a son, but you might win more as well." ■

Clockwise from top left: Caño Negro Wildlife Refuge; Tortuguero Park Canal Cruise; Toucan; Rainforest; Manuel Antonio Park

Costa Rica Natural Paradise 10 Days $995

Call Now for Choice Dates

Caravan makes it so easy-and so affordable- for you to explore the magnificent rainforests, beaches and volcanoes of Costa Rica. Your Costa Rica tour includes all meals, all activities, all hotels, a great itinerary, all airport transfers, all transportation and excursions within Costa Rica.

A professional tour director accompanies you for the entire tour. With naturalist guides, see exotic birds and wildlife, hike in jungle rainforests, view volcanoes, soak in hot springs and cruise biological reserves.

Choose Your Tour—Only $995

8 days	Canadian Rockies
8 days	Nova Scotia & Prince Edward Isl.
8 days	Grand Canyon, Zion, Bryce
8 days	Mount Rushmore, Yellowstone
8 days	California Coast and Yosemite
8 days	New England Fall Colors
8 days	Mexico's Copper Canyon Train
8 days	Mexico's Ancient Cultures
11 days	Guatemala with Tikal and Copan
10 days	Costa Rica Natural Paradise

"Caravan is very reasonably priced" — New York Times

#1 in Value. Fully Escorted Tours Since 1952

Caravan.

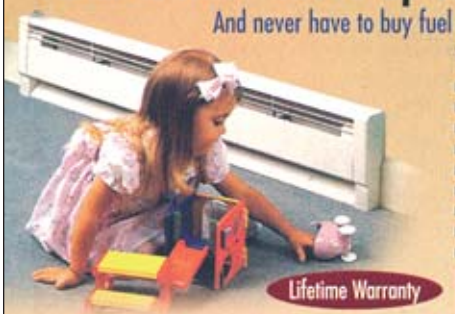
Call Now FREE 28 Page Info Guide

1-800-Caravan Caravan.com

MOST DEPARTURES $995 PER PERSON, DOUBLE OCCUPANCY, PLUS TAXES AND FEES.

Families Have Saved Up To 50% On Heating Costs

And never have to buy fuel — wood, oil, gas, kerosene — ever again!



Lifetime Warranty

Your Benefits with Hydro-Sil:

- Slash heating cost with Hydro-Sil technology
- Lifetime warranty. No service contracts
- Safe, complete peace of mind
- Clean, no fumes, environmentally safe
- U.L. listed
- Preassembled — ready to use
- No furnaces, ducts, or chimneys
- Portable (110V) or permanent (220V)
- Whole house heating or single room

Proportional energy saving thermostat!

Hydro-Sil is a duct-free/furnace-free high performance room-by-room heating system that can save you hundreds of dollars in home heating costs by replacing old and inefficient heating. It can replace or supplement your electric heat, gas or oil furnace and woodstoves.

Hydro-Sil represents economy in heating: inside the heater case is a sealed copper tube filled with a harmless silicone fluid. It's permanent. You'll never run out. Hydro-Sil "Energy Smart" thermostat controls a variable watt hydro element that is only being supplied a proportional amount of power on an as-needed basis. The silicone liquid is quickly heated, and with its heat retention qualities, continues to heat after the Hydro element shuts off. Hydro-Sil's digital proportional control technology greatly increases energy savings and comfort.

220 VOLT PERMANENT	Approx. Area to Heat	Discount Price	Quantity
8' 2000 watts	250-300 s.f.	$309	
6' 1500 watts	180-250 s.f.	$279	
5' 1250 watts	130-180 s.f.	$259	
4' 1000 watts	100-130 s.f.	$229	
3' 750 watts	75-100 s.f.	$189	
2' 500 watts	50-75 s.f.	$169	
Thermostats	Call for options & exact heater needed		

110 VOLT PORTABLES (Thermostat Included)	Discount Price	Quantity
5' Hydro-Max 750-1500 watts	$229	
4' Convector – Dual watt	$179	
3' 750 watts – Silicone	$179	
$18.00 shipping per heater	$	
Total Amount	$	

Order today or contact us for more info.

Check ■ MasterCard ■ Visa

1-800-627-9276

www.hydrosil.com

Hydro-Sil, P.O. Box, 662, Fort Mill, SC 29715

Name_____

Address_____

City_____ St_____ Zip_____

Phone_____

MasterCard or Visa Account Information:

Acct #_____

Expiration Date_____

© 2007 SCIENTIFIC AMERICAN, INC.

Page Intentionally Blank

SCIENTIFIC AMERICAN Digital

Heating Up

By Mark Fischetti

More than 30 states have passed or are considering "renewable energy portfolio standards" that require utility companies to generate some portion of their electricity from renewable sources. Geothermal power plants, which tap hot subterranean water or steam, are high on many lists.

Most utilities have not pursued geothermal energy primarily because up-front costs, including exploratory drilling, can be high. (Geothermal taps deep reservoirs, not groundwater, which collects much closer to the surface.) But once in operation, the plants consume no fuel and create few if any emissions. "When looking at the true costs over a plant's lifetime, geothermal is on par with or better than a coal plant," the least expensive conventional option, notes Gerald Nix, geothermal technology manager at the National Renewable Energy Laboratory in Golden, Colo.

Furthermore, an extensive study recently released by the Massachusetts Institute of Technology shows that the heat available underground is surprisingly plentiful nationwide. "Geothermal has been dramatically underutilized," Nix concurs.

Several different power plant designs have been operating for years in select places, depending on how hot the water is. So-called flash plants are most common [*see main illustration on opposite page*]. In the future, however, "binary plants will predominate," Nix says. Binary designs, which use the water to vaporize a second fluid, can operate with lower-temperature water, making geothermal viable in more places.

Some critics worry that the reservoirs could slowly be depleted because some of the water can be lost during flash conversion to steam as well as when that steam is subsequently cooled. But the inner earth will naturally replenish reservoirs unless water is drawn too fast. And binary plants send virtually all the extracted water back into the reservoir, although they may cost more to run than flash plants.

Looking ahead, utilities could even use "enhanced" recovery methods to get steam from hot, dry rock [*see illustration on this page*]. Homeowners, too, could tap their own backyards. At a mere 10 feet deep, the ground remains a constant 50 to 60 degrees Fahrenheit year-round. Fluid-filled pipes laid there can feed a home heat pump that provides cooling in summer and heating in winter. "When you are building a new house, the heat-pump system does tend to cost more than a conventional furnace," Nix notes. But it doesn't consume fuel, other than a small amount of electricity. "In four or five years you break even and start saving money," he says.

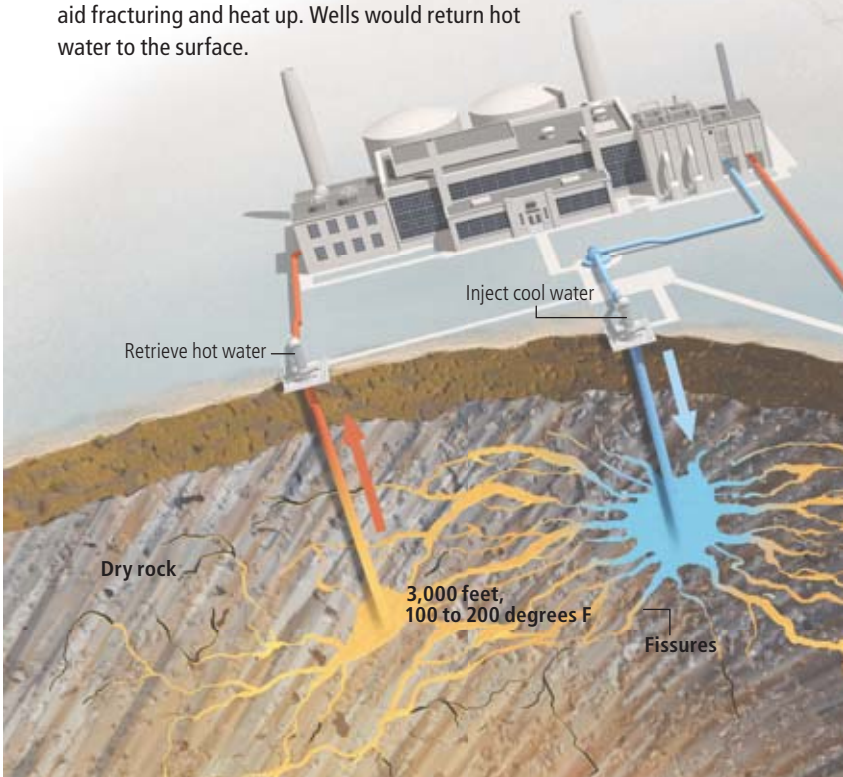
DID YOU KNOW ...

LARGE BUT SMALL: Twenty-four countries have geothermal power plants, with an overall combined electric generating capacity of 8,900 megawatts. The U.S. boasts the largest share, at 2,850 MW (2,490 MW in California), although the total accounts for only 0.36 percent of the nation's electricity supply. Since 2000 geothermal generation has tripled in France, Kenya and Russia.

OLD FAITHFUL: The Geysers, 72 miles north of San Francisco, is the world's largest geothermal power complex. Begun in 1960, it currently operates 21 plants with a combined capacity of 750 megawatts. The city of Santa Rosa now conveniently pipes treated wastewater there; it is injected into the underground reservoir, prolonging the reservoir's useful lifetime.

NOT ALL CLEAN: Some deep waters contain condensed gases such as carbon dioxide or hydrogen sulfide and minerals such as zinc. These impurities must be captured or treated at flash plants. Binary plants send the compounds back to the reservoir.

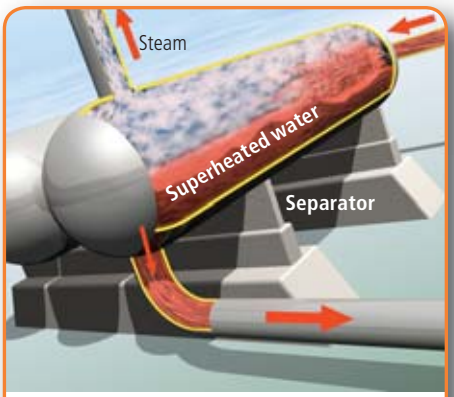
ENHANCED PLANT of the future would inject cool, pressurized water into dry fissures, where it would aid fracturing and heat up. Wells would return hot water to the surface.



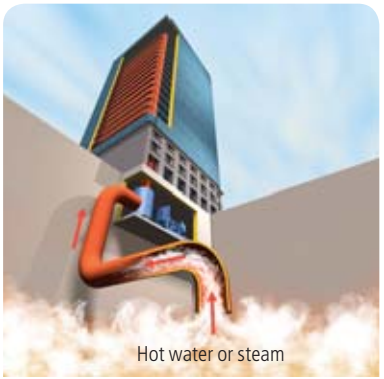
GEOTHERMAL POWER PLANTS tap into hot underground water. Flash designs are most common, but binary schemes may prevail in the future. The plants create high-pressure steam or vapor that turns a turbine. Leftover water is injected back into the reservoir.

NICK ROTONDO

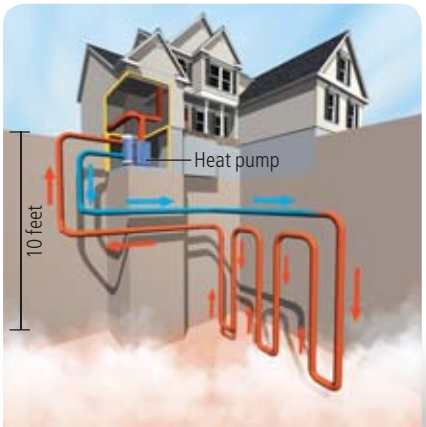
© 2007 SCIENTIFIC AMERICAN, INC.



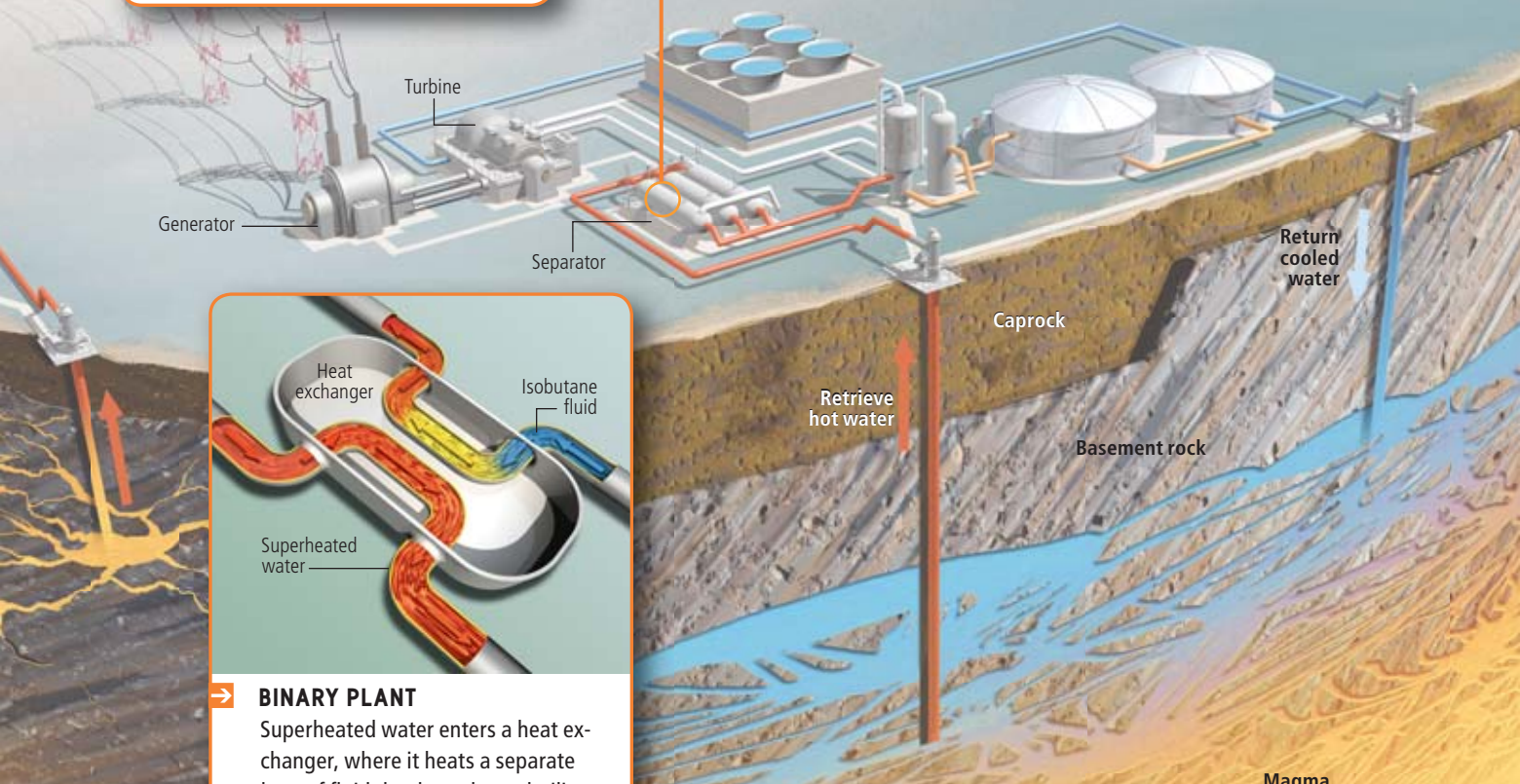
FLASH PLANT
Superheated water under naturally high pressure enters a separator at lower pressure. The pressure drop causes some of the fluid to vaporize, or flash, into steam. Remaining water enters a second chamber at even lower pressure, where it flash-boils, too.



BUILDINGS can be heated directly with piped-in hot water or steam that rises close to the surface in select spots.



HOUSE heat pump circulates fluid through shallow ground pipes or wells, which stay at 50 to 60 degrees Fahrenheit year-round. In summer the relatively cool fluid draws heat from inside the house. In winter the relatively warm fluid carries heat into the house.



BINARY PLANT
Superheated water enters a heat exchanger, where it heats a separate loop of fluid that has a lower boiling point, such as isobutane. That fluid boils, creating high-pressure vapor.

SEND TOPIC IDEAS *to*
workingknowledge@SciAm.com

© 2007 SCIENTIFIC AMERICAN, INC.

Rapturous Sociability ▪ Armageddon Avoided ▪ The Allure of Venus

BY MICHELLE PRESS

SHYNESS: HOW NORMAL BEHAVIOR BECAME A SICKNESS

by Christopher Lane. Yale University Press, 2007 ($27.50)

Would Henry David Thoreau and Emily Dickinson be given drugs today? In the 1980s a small group of leading psychiatrists revised the profession's diagnostic manual, called the *DSM* for short, adding social anxiety disorder—aka shyness—and dozens of other new conditions. Christopher Lane, Miller Research Professor at Northwestern University, uses previously secret documents, many from the American Psychiatric Association archives, to support his argument that these decisions were marked by carelessness, pervasive influence from the pharmaceutical industry, academic politics and personal ambition. Lane shows how drug companies seized on the newly minted disorders to sell millions of dollars' worth of psychotropic drugs. Some have dangerous side effects; some were already developed—treatments looking for a disease. The next revision of the *DSM* is already under way, and Lane warns that without drastic reform many more common behaviors—excessive shopping, poorly controlled anger, defiance—can become pathologies for which drugs are already on tap.

ARSENALS OF FOLLY: THE MAKING OF THE NUCLEAR ARMS RACE

by Richard Rhodes. Alfred A. Knopf, 2007 ($28.95)

As much about the unmaking as the making of the superpower arms race, Richard Rhodes's latest book alludes in its subtitle to his Pulitzer Prize–winning *The Making of the Atomic Bomb.* He begins by tracing the convictions that shaped two men's determination to end the threat of nuclear annihilation. Mikhail Gorbachev,

born into a peasant family in 1931, lived through Stalin's "Great Terror," when the dictator deliberately starved millions of farmers to force them onto collective farms. The destruction at Chernobyl—on the order of one-third the power of the *smallest* nuclear explosive—crystallized his mission. "Global nuclear war can no longer be the continuation of rational politics, as it would bring the end of all life," Gorbachev told the Politburo in 1986.

Ronald Reagan was molded into the unlikely opposite side of this equation by religion and movies. He connected nuclear war with Armageddon and believed he was predestined to do away with nuclear weapons. After seeing the 1983 film *The Day After,* in which Lawrence, Kan., is wiped out in a nuclear war with Russia, Reagan wrote in his diary that it "left me greatly depressed.... we have to do all we can ... to see that there is never a nuclear war." The book culminates in a highly detailed and gripping account of the meeting of the two leaders in Reykjavik in 1986, where, against all odds, they managed to halt the buildup of nuclear arsenals. "Leaders who make history are often provincials. Provincials attempt what sophisticates consider naïve," Rhodes writes, before turning to the world of the present—and a chilling conclusion.

EXCERPT

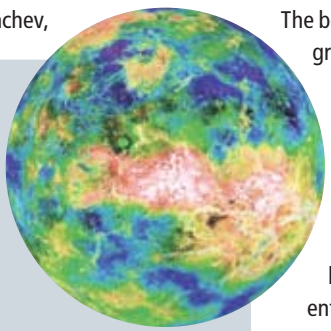
LIVES OF THE PLANETS: A NATURAL HISTORY OF THE SOLAR SYSTEM

by Richard Corfield. Basic Books, 2007 ($30)

In his wonderfully written exploration of the solar system, Richard Corfield, a planetary scientist at the Open University in England, describes the fascination with Venus, "the greenhouse in the sky," in the early days of space probes:

"On December 14, 1962, [the U.S. spacecraft] *Mariner 2* grazed Venus, skimming past the planet at a distance of only 21,000 miles.... The results were unequivocal.... The surface of Venus is ... as hot as the interior of a self-cleaning oven ... no global oceans, no swamps, no giant tree ferns, no enormous insects, and no amphibian-like creatures crawling their way toward sentience.

"One immediate effect of the news from *Mariner 2* was that America lost all interest in Venus.... In startling contrast, the Soviet focus on Venus intensified ... the message from the Central Committee was clear: the Soviet space establishment was to focus the attention of its nascent unmanned space program on this nearby, bright planet that glimmers so temptingly in the evening sky, with the goal of landing a probe on it. Such an order was nothing if not audacious for, at this time in the early 1960s, no one had even landed a probe on the moon."



VENUS as imaged by NASA's *Magellan* spacecraft; false color indicates elevation.

NEW BOOKS ON ENERGY AND THE ENVIRONMENT

1 **This Cold House: The Simple Science of Energy Efficiency**
by Colin Smith. Johns Hopkins University Press, 2007 ($25)

2 **Break Through: From the Death of Environmentalism to the Politics of Possibility**
by Ted Nordhaus and Michael Shellenberger. Houghton Mifflin, 2007 ($25)

3 **Green to Gold: How Smart Companies Use Environmental Strategy to Innovate, Create Value, and Build a Competitive Advantage**
by Daniel C. Esty and Andrew Winston. Yale University Press, 2007 ($27.50)

4 **Lights Out: The Electricity Crisis, the Global Economy, and What It Means to You**
by Jason Makansi. Wiley, 2007 ($27.95)

NASA/JPL-CALTECH

© 2007 SCIENTIFIC AMERICAN, INC.

Q *What is a "fictitious force"?*

—E. Lopez, Los Angeles

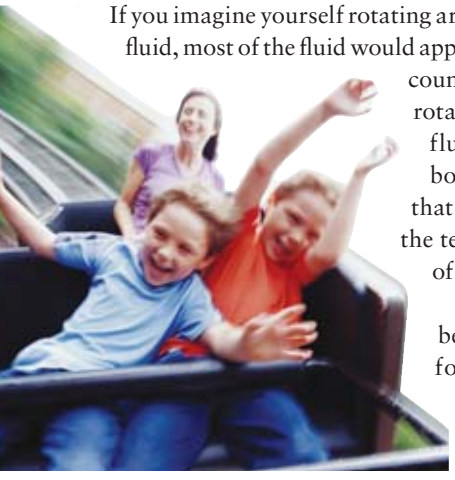
California Institute of Technology theoretical physicist and 2004 Nobel laureate **David Politzer** replies:

The forces you feel in a moving car—those that push you back into your seat when the driver steps on the gas or throw you sideways when the car makes sharp turns—are everyday examples of fictitious forces. In general, these influences arise because the natural frame of reference for a given situation is itself accelerating.

The term "fictitious force" has a precise meaning within Newtonian mechanics—in fact, it is always proportional to the mass of the object on which it acts.

An elegant example of these types of apparent influences is the Coriolis force, which is responsible for the stately precession (or circular rotation) of a carefully suspended pendulum's plane of swing. If such a pendulum were directly above the North Pole, it would appear to an earthly observer to rotate 360 degrees every day. If you viewed this pendulum from a stationary point in outer space, however, it would appear to swing in a single, fixed plane while the Earth turned under it. From the outer-space perspective, there is no sideways force (that is, perpendicular to the plane of swing) deflecting the pendulum's sway. That is why the somewhat pejorative term "fictitious" is attached to this force. Similarly, in the car, no real force pushes you back into your seat, your senses notwithstanding.

Tea leaves offer a charming demonstration of a consequence of the Coriolis force. If a few leaves are present in a stirred cup of tea, they end up in a central pile at the bottom of the cup (and not along the edge, as one might expect, as a result of the also fictitious centrifugal force). If you imagine yourself rotating around in sync with the stirred fluid, most of the fluid would appear to be at rest while the cup counterrotates around you. That rotating cup drags some adjacent fluid along with it. Near the bottom, the Coriolis force on that dragged fluid pushes it—and the tea leaves—toward the center of the cup.



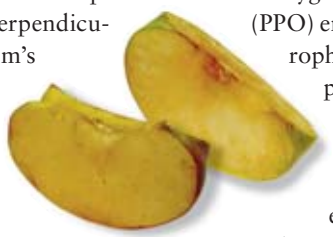
With general relativity, Albert Einstein managed to blur forever the distinction between real and fictitious forces. General relativity is his theory of gravity—certainly the paradigmatic example of a "real" force. The cornerstone of Einstein's theory, however, is the proposition that gravity is itself a fictitious force (or, rather, that it is indistinguishable from a fictitious force). Now, some 90 years later, we have innumerable and daily confirmations that his theory appears to be correct.

Q *Why do* ***apple slices*** *turn brown after being cut?*

—A. Suraya, Guntur, India

Lynne McLandsborough, a professor of food science at the University of Massachusetts Amherst, explains:

When an apple is cut or bruised, oxygen introduced into its injured tissue reacts with compounds there, leading to browning. When oxygen is present in the tissue's cells, polyphenol oxidase (PPO) enzymes in the chloroplasts (sites that harbor chlorophyll and conduct photosynthesis) rapidly oxidize phenolic compounds naturally present in apple tissues to o-quinones. These colorless compounds then self-assemble into polymers or react with amino acids or proteins, forming dark-pigmented secondary products in the fruit tissue.



This enzymatic browning can be prevented by either reducing PPO oxidation or lowering the amount of substrate to which the enzyme can bind. Coating freshly cut apples in sugar or syrup can reduce oxygen diffusion and thus slow the browning reaction. Lemon or pineapple juices, both of which contain antioxidants, also can delay discoloration. (In addition, both fruit juices are acidic, and the lower pH that they bring about reduces PPO activity.) And although it changes the fruit's texture, heating can inactivate PPO enzymes—blanching apples in boiling water for four to five minutes will nearly eliminate PPO activity.

Enzymatic browning is not unique to apples. PPO is present in nearly all plant tissues and also exists in bacteria, animals and fungi. In fact, browning from PPO is not always unwanted: the familiar shades of tea, coffee and cocoa come from PPO enzymatic browning during product processing. ■

HAVE A QUESTION?... *Send it to experts@SciAm.com or go to* **www.SciAm.com/asktheexperts**

GETTY IMAGES (*roller coaster*); EMILY ALAMSON (*apples*)

© 2007 SCIENTIFIC AMERICAN, INC.

Do Helmets Attract Cars to Cyclists?

BY NIKHIL SWAMINATHAN

You may never forget how to ride a bicycle, but should you forget your helmet when you hop on your two-wheeler?

About a year ago a psychologist at the University of Bath in England announced the results of a study in which he played both researcher and guinea pig. An avid cyclist, Ian Walker had heard several complaints from fellow riders that wearing a helmet seemed to result in bike riders receiving far less room to maneuver—effectively increasing the chances of an accident. So Walker attached ultrasonic sensors to his bike and rode around Bath, allowing 2,300 vehicles to overtake him while he was either helmeted or bareheaded. In the process, he was actually contacted by a truck and a bus, both while helmeted—though, miraculously, he did not fall off his bicycle either time.

Cancels Out Its Protection

Walker's findings, published in the March issue of *Accident Analysis & Prevention,* state that when he wore a helmet drivers typically drove an average of 3.35 inches closer to his bike than when his head was not covered. But if he wore a wig of long, brown locks in lieu of a helmet—appearing to be a woman from behind—he was granted 2.2 inches more room to ride.

"The implication," Walker says, "is that any protection helmets give is canceled out by other mechanisms, such as riders possibly taking more risks and/or changes in how other road users behave toward cyclists." He cannot explain the extra leeway granted to him when he pretended to be a woman, although he speculates that drivers might perceive members of the fairer sex as less capable riders, frailer or just less frequent bikers than men.

Still Enough Room to Ride, Right?

Randy Swart, founder of the Bicycle Helmet Safety Institute, asserts that studies such as Walker's run the risk of misleading cyclists as to the effectiveness of helmets. "The cars were giving him, on average, a very wide passing clearance already," he explains, noting that most vehicles typically stayed well over three feet from Walker's bike, rendering the 3.35-inch discrepancy insignificant.

Walker reanalyzed his data recently to address this line of reasoning. "I assessed the number of vehicles coming within one meter [roughly 3.3 feet] of the rider, on the principle that these are the ones that pose a risk," he says. "There were 23 percent more vehicles within this one-meter danger zone when a helmet was worn, suggesting a real risk."

Before Walker's paper came out, Dorothy Robinson of the University of New England in Armidale, Australia, published a review article in the *British Medical Journal* about what happened after regions in Australia, New Zealand and Canada introduced legislation that spurred a more than 40 percent increase in bicycle helmet use among their populaces. The rates of bike accidents involving head injuries were on the decline before the newly instituted laws, and she found the new mandates did not result in a sudden drop-off in the percentage of cyclists sustaining head injuries.

Should You Protect Your Noggin?

Coincidentally, around the same time that Walker announced his results, New York City released a report on bicycle deaths and injuries strongly suggesting that helmets *do* reduce head injuries: 225 cyclists died between 1996 and 2005 on city streets; 97 percent of them were not wearing helmets. Of these deaths, 58 percent are known to have involved head injury, but the actual number could be as high as 80 percent. Comparing the helmet to a seat belt in a car, Swart says, "When you do have that crash, you had better have it on."

In agreement with Swart are officials at the Centers for Disease Control and Prevention, who point out that helmets are considered an effective way to help prevent head injuries from bicycle mishaps. Walker will not make specific recommendations to other cyclists about whether to wear headgear. But he does urge people to read the research. And watch out for cars. ■

AGE FOTOSTOCK

Nikhil Swaminathan is a reporter for www.SciAm.com

MORE ONLINE... *Find more investigations into popular myths at* **www.SciAm.com/factorfiction**

To suggest a topic, write to factorfiction@SciAm.com

© 2007 SCIENTIFIC AMERICAN, INC.